

# Space–Time Codes Achieving the DMD Tradeoff of the MIMO-ARQ Channel

Sameer A. Pawar, K. Raj Kumar, *Student Member, IEEE*, Petros Elia, P. Vijay Kumar, *Fellow, IEEE*, and B. A. Sethuraman

**Abstract**—For the quasi-static, Rayleigh-fading multiple-input multiple-output (MIMO) channel with  $n_t$  transmit and  $n_r$  receive antennas, Zheng and Tse showed that there exists a fundamental tradeoff between diversity and spatial-multiplexing gains, referred to as the diversity–multiplexing gain (D-MG) tradeoff. Subsequently, El Gamal, Caire, and Damen considered signaling across the same channel using an  $L$ -round automatic retransmission request (ARQ) protocol that assumes the presence of a noiseless feedback channel capable of conveying one bit of information per use of the feedback channel. They showed that given a fixed number  $L$  of ARQ rounds and no power control, there is a tradeoff between diversity and multiplexing gains, termed the diversity–multiplexing–delay (DMD) tradeoff. This tradeoff indicates that the diversity gain under the ARQ scheme for a particular information rate is considerably larger than that obtainable in the absence of feedback.

In this paper, a set of sufficient conditions under which a space–time (ST) code will achieve the DMD tradeoff is presented. This is followed by two classes of explicit constructions of ST codes which meet these conditions. Constructions belonging to the first class achieve minimum delay and apply to a broad class of fading channels whenever  $n_r \geq n_t$  and either  $L|n_t$  or  $n_t|L$ . The second class of constructions do not achieve minimum delay, but do achieve the DMD tradeoff of the fading channel for all statistical descriptions of the channel and for all values of the parameters  $n_r$ ,  $n_t$ ,  $L$ .

**Index Terms**—Automatic retransmission request (ARQ), cyclic division algebra, diversity–multiplexing–delay (DMD) tradeoff, diversity–multiplexing gain (D-MG) tradeoff, explicit construction, multiple-input multiple-output (MIMO) feedback, space–time (ST) codes.

Manuscript received November 01, 2005; revised December 01, 2007. Current version published June 24, 2009. This work was supported in part by the DRDO-IISc Program on Advanced Research in Mathematical Engineering and in part by NSF-ITR CCR-0326628. The material in this paper was presented in part at the IEEE International Symposium on Information Theory, Adelaide, Australia, September 2005. Much of this work was carried out while S. A. Pawar and K. R. Kumar were with the Indian Institute of Science, Bangalore, India, and P. Elia was with the University of Southern California, Los Angeles.

S. A. Pawar is with the Department of Electrical Engineering and Computer Science, University of California Berkeley, Berkeley, CA 94709 USA (e-mail: sameerpawar@berkeley.edu).

K. R. Kumar is with the Department of Electrical Engineering–Systems, University of Southern California, Los Angeles, CA 90089 USA (e-mail: rkkrishn@usc.edu).

P. V. Kumar is with the Department of Electrical Communication Engineering, Indian Institute of Science, Bangalore, 560 012, India, on leave from the University of Southern California, Los Angeles, CA 90089 USA (e-mail: vijay@ece.iisc.ernet.in).

P. Elia is with EURECOM, Sophia-Antipolis, France (e-mail: Petros.Elia@eurecom.fr).

B. A. Sethuraman is with the Department of Mathematics, California State University, Northridge, CA 91330 USA (e-mail: al.sethuraman@csun.edu).

Communicated by H. Boche, Associate Editor for Communications.

Color versions of Figures 1 and 6 in this paper are available online at <http://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/TIT.2009.2021332

## I. INTRODUCTION

IN the quasi-static or equivalently, block-fading space–time (ST) channel with quasi-static interval  $T$ ,  $n_t$  transmit, and  $n_r$  receive antennas, the  $(n_r \times T)$  received signal matrix  $Y$  is given by

$$Y = \theta H X + W \quad (1)$$

where  $\theta X$  is the  $(n_t \times T)$  transmitted matrix drawn from an ST code  $\mathcal{X}_{n_t \times T}$ ,  $H$  is the  $(n_r \times n_t)$  channel matrix, and  $W$  is the  $(n_r \times T)$  noise matrix. The entries of  $W$  are assumed to be independent and identically distributed (i.i.d.), circularly symmetric complex Gaussian  $\mathcal{CN}(0, 1)$  random variables. The entries of  $X$  are drawn from a constellation whose size scales with the signal-to-noise ratio (SNR) and the scaling factor  $\theta$  is chosen to ensure the energy constraint

$$\mathbb{E}(\|\theta X\|_F^2) \leq T \text{SNR}. \quad (2)$$

### A. ARQ Signaling

As in [1], our interest here is in signaling across the channel in (1) using an automatic retransmission request (ARQ) protocol. Under this protocol, each message symbol from the source is associated with a unique block  $\theta[X_1 X_2 \cdots X_L]$  of matrices, each  $X_i \in \mathbb{C}^{n_t \times T}$  in such a way that it is possible to uniquely decode the message symbol given  $\theta[X_1 X_2 \cdots X_l]$  for any  $1 \leq l \leq L$ , in particular, given just  $\theta X_1$ . The scalar  $\theta$  is chosen to ensure the energy constraint

$$\mathbb{E}[\theta^2 \|X_l\|_F^2] \leq T \text{SNR}, \quad 1 \leq l \leq L \quad (3)$$

is satisfied.

**Bank of Codes:** Let  $\mathcal{X}_{\text{ARQ}}$  denote the ARQ ST code, i.e., the collection of matrices  $\theta[X_1 X_2 \cdots X_L]$  corresponding to all possible message symbols. We will use  $\mathcal{X}_{\text{ARQ},l}$  to denote the ST code truncated to  $l$  rounds, i.e., the collection of  $(n_t \times lT)$  matrices  $\theta[X_1 X_2 \cdots X_l]$  corresponding to all possible message symbols. We will refer to  $\mathcal{X}_{\text{ARQ},l}$ ,  $1 \leq l \leq L$  as the  $l$ th round ST code and to the specific codes  $\mathcal{X}_{\text{ARQ},1}$  and  $\mathcal{X}_{\text{ARQ},L}$  as the *single-round* and *full-length* ST codes, respectively.

With each code  $\mathcal{X}_{\text{ARQ},l}$ , we associate a decoder  $\mathcal{D}_l$ . While each of the decoders  $\mathcal{D}_l$ ,  $1 \leq l \leq (L-1)$  is permitted to decline to decode, the decoder  $\mathcal{D}_L$  in contrast, is a maximum-likelihood (ML) decoder and will always decode.

**ARQ Signaling:** The distinction between the ARQ ST code  $\mathcal{X}_{\text{ARQ}}$  and the full-length ST code  $\mathcal{X}_{\text{ARQ},L}$  is the manner in which the sequence of matrices  $\{\theta X_i\}$  is used. With

the full-length code, one always transmits the entire matrix  $\theta[X_1 X_2 \cdots X_L]$ . With the ARQ code, one first transmits the matrix  $\theta X_1$ . The presence of a noiseless feedback channel capable of conveying one bit (ACK or NACK) of information per ARQ round is assumed. Also assumed is the presence of an infinite buffer at the transmitter end. At the receiver end, the receiver applies the decoder  $\mathcal{D}_1$  to the corresponding  $(n_r \times T)$  received matrix  $Y_1$ . If decoder  $\mathcal{D}_1$  is able to decode the underlying message symbol from  $Y_1$ , then an ACK is passed on to the transmitter, otherwise, a NACK is sent. Upon receipt of an ACK, the transmitter moves on to transmit the next message symbol. Upon receipt of a NACK, however, the transmitter proceeds to transmit  $\theta X_2$ . Upon receipt of the corresponding received signal matrix  $Y_2$ , the receiver once again attempts to decode the message symbol, this time applying decoder  $\mathcal{D}_2$  to the concatenated signal  $[Y_1 Y_2]$ . This process is continued until the transmitter receives an ACK. Note that since the decoder  $\mathcal{D}_L$  is an ML decoder, the transmitter will receive an implicit ACK after the  $L$ th round and hence the ARQ signaling will never run beyond  $L$  rounds.

To simplify notation, we will abbreviate and write

$$\begin{aligned} \mathbf{X}_l &= [X_1 X_2 \cdots X_l] \\ \mathbf{Y}_l &= [Y_1 Y_2 \cdots Y_l] \\ \mathbf{W}_l &= [W_1 W_2 \cdots W_l] \end{aligned}$$

for  $1 \leq l \leq L$ . These are related by

$$\mathbf{Y}_l = \theta[H_1 X_1 \ H_2 X_2 \ \cdots \ H_l X_l] + \mathbf{W}_l. \quad (4)$$

Note that by our assumption above relating to unique decodability, all the codes  $\mathcal{X}_{\text{ARQ},l}$  have the same cardinality, i.e.,

$$|\mathcal{X}_{\text{ARQ},l}| = |\mathcal{X}_{\text{ARQ}}|, \quad 1 \leq l \leq L. \quad (5)$$

In this paper, we will assume the long-term static channel model introduced in [1], which assumes that the matrices  $H_i$  encountered over the  $L$  rounds are identical. Under this assumption, (4) takes on the simpler form

$$\mathbf{Y}_l = \theta H \mathbf{X}_l + \mathbf{W}_l, \quad 1 \leq l \leq L. \quad (6)$$

As in [1], the matrices  $H$  associated with the transmission of different message symbols will, however, be assumed to be statistically independent.

### B. Rate and Reliability

Let  $R$  denote the rate of the ARQ scheme, i.e., the average number of bits transmitted per channel use. Let  $r$  denote the normalized rate given by

$$R = r \log_2(\text{SNR}).$$

We will refer to  $r$  as the spatial-multiplexing gain [2], [3]. At times, we will also refer to  $r$  as the normalized (with respect to SNR) rate. The relationship between the multiplexing gain  $r$  and the size  $|\mathcal{X}_{\text{ARQ}}|$  of the ARQ ST code  $\mathcal{X}_{\text{ARQ}}$  will be established in Section II. Note that our notation here differs somewhat from that in [1], in which the rate  $R$  is termed the ‘‘throughput’’ and denoted by  $\eta$ .

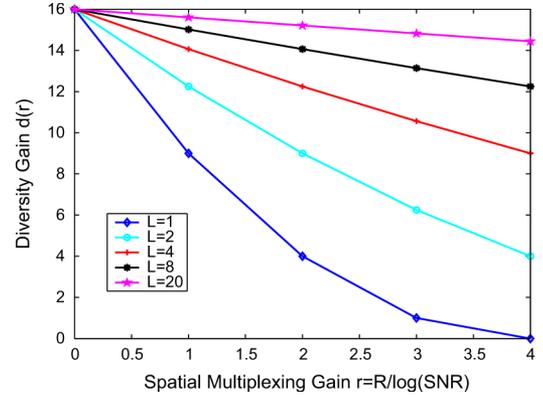


Fig. 1. The D-MG tradeoff of the MIMO-ARQ channel as a function of the number  $L$  of ARQ rounds. Here  $n_t = n_r = 4$ .

For a given normalized rate  $r$ , let the probability of a message symbol being decoded incorrectly at the receiver be denoted by  $P_e(r)$ . We define the diversity gain  $d(r)$  as

$$d(r) = - \lim_{\text{SNR} \rightarrow \infty} \frac{\log P_e(r)}{\log(\text{SNR})}.$$

This implies that for large SNR,  $P_e(r)$  decays as  $\text{SNR}^{-d(r)}$ . In the exponential-equality notation of [3], we will denote this by

$$P_e(r) \doteq \text{SNR}^{-d(r)}.$$

Then under this setting, El Gamal *et al.* [1] showed that for channels with Rayleigh fading, the maximum possible value  $d_{\text{Rayleigh}}^*(r)$  of the diversity gain at each value of spatial-multiplexing gain  $r$  is given by

$$d_{\text{Rayleigh}}^*(r) = f\left(\frac{r}{L}\right)$$

for  $r$  in the range  $0 \leq r \leq \min\{n_t, n_r\}$  and 0 outside the range, where  $f(k)$  is the piecewise-linear function connecting the points  $(k, (n_t - k)(n_r - k))$  for integral values of  $0 \leq k \leq \min\{n_t, n_r\}$ . The function  $f(\cdot)$  represents the tradeoff between the diversity and spatial-multiplexing gains, developed by Zheng and Tse [3] for the Rayleigh-fading channel in the absence of feedback, referred to as the diversity–multiplexing gain (D-MG) tradeoff. Thus,  $f(r/L)$  represents the DMD for a given number  $L$  of ARQ rounds and is shown plotted in Fig. 1 for the case of four transmit and four receive antennas and for varying  $L$ . The bottom plot corresponding to  $L = 1$  (and hence, no ARQ) represents the original Zheng–Tse D-MG tradeoff.

An impressive amount of recent research has culminated in the construction of explicit ST codes that meet the optimal D-MG tradeoff for any  $n_t, n_r$ , for arbitrary fading channels, see [4]–[12] and references therein for details.

Returning to the ARQ setting, the tradeoff between diversity and multiplexing gains is clearly, from Section I-B, a function of the parameter  $L$  that indicates the maximum number of ARQ rounds and is also indicative of the maximum delay  $LT$  of channel uses before one is guaranteed to be in a position to decode the transmitted message. Accordingly, the authors of [1] term the tradeoff in the ARQ case, the diversity–multiplexing–delay (DMD) tradeoff. It is shown in [1], that an adaptation of the LAST codes [7] for the ARQ, Rayleigh-fading

channel known as IR (incremental redundancy) LAST codes, are optimal with respect to the corresponding DMD tradeoff.

### C. Results

In this paper, a set of sufficient conditions under which an ST code will achieve the DMD tradeoff is presented. This is followed by two classes of explicit constructions of ST codes which meet these conditions. Constructions belonging to the first class achieve minimum delay and apply to a broad class of fading channels which we label as the class of *regular fading channels* and which are defined below. These constructions are applicable whenever  $n_r \geq n_t$  and either  $L|n_t$  or  $n_t|L$ . The second class of constructions do not achieve minimum delay, but do achieve the DMD of the fading channel for all statistical descriptions of the channel and for all values of the parameters  $n_r, n_t, L$ . The second class of constructions may hence be said to possess the approximate universality property [8].

We will say that a fading channel is *regular* if the associated channel matrix  $H$  is such that the density function  $p_{\lambda_{\min}}(\lambda)$  of the smallest eigenvalue  $\lambda_{\min}$  of  $H^\dagger H$  is finite and well-behaved near zero, i.e., for  $\epsilon$  sufficiently small, we have

$$\int_0^\epsilon p_{\lambda_{\min}}(\lambda) d\lambda \approx \epsilon p_{\lambda_{\min}}(0).$$

Channels with Rayleigh fading, i.e., channels whose  $H$  matrix has components that are i.i.d., circularly symmetric complex Gaussian  $\mathcal{CN}(0, 1)$  random variables, fall into this class, as can be verified from the expression

$$p_{\lambda_{\min}}(\lambda) = c\lambda^{n-m} e^{-\lambda m/2} P_{m,n}(\lambda)$$

for the density function in the Rayleigh case, given in [17], where  $n = \max\{n_r, n_t\}$ ,  $m = \min\{n_r, n_t\}$ ,  $c$  is a normalizing constant, and  $P_{m,n}(\lambda)$  is a polynomial of degree  $(n-m)(m-1)$ .

The sufficiency condition for achieving the DMD tradeoff is presented in Section II. The minimum-delay constructions are described in Section III, while those that are approximately universal may be found discussed in Section IV. Simulation results are presented in the final section, Section V. Most proofs are relegated to the Appendices.

## II. SUFFICIENT CONDITION FOR ACHIEVING THE DMD TRADEOFF

We begin with a sufficient condition for an ST code to achieve the D-MG tradeoff given  $n_t, n_r$  and quasi-static interval  $T$ .

### A. A Sufficient Condition for Achieving the D-MG Tradeoff

*Theorem 1:* Consider an  $(n_t \times T)$  ST code  $\mathcal{X}$  with  $T \geq n_t$  of size  $|\mathcal{X}| = \text{SNR}^{rT}$ . Let  $\theta \Delta X$  denote the difference of any two code-matrices drawn from  $\mathcal{X}$ . Define  $\min_{\Delta X} \det(\theta^2 \Delta X \Delta X^\dagger) := \text{SNR}^\delta$ . If

$$\delta = n_t - r$$

then  $\mathcal{X}$  is optimal with respect to the D-MG tradeoff at multiplexing gain  $r$  for any number of receive antennas and over any fading channel.

A proof of the theorem for the special case of the Rayleigh fading may be found in [10]. The proof in the case of the general fading channel, appears in [8].

### B. A Sufficient Condition for Achieving the DMD Tradeoff

We begin with some definitions. At the conclusion of the  $l$ th round of transmission, the receiver examines the matrix  $\mathbf{Y}_l$  using decoder  $\mathcal{D}_l$  and then makes a decision whether to send an ACK or a NACK. Let  $\mathcal{A}_l (\bar{\mathcal{A}}_l)$  be the event that the received matrix  $\mathbf{Y}_l$  is such that upon receipt of  $\mathbf{Y}_l$ , the receiver will send an ACK (NACK).

Next, let  $R_l$  be the rate in bits per channel use associated with the  $l$ th ARQ code  $\mathcal{X}_{\text{ARQ},l}$ , i.e.,

$$2^{lR_l T} = |\mathcal{X}_{\text{ARQ},l}| = |\mathcal{X}_{\text{ARQ}}|, \quad 1 \leq l \leq L$$

from (5).<sup>1</sup> Let  $r_l$  denote the corresponding normalized rate given by

$$R_l = r_l \log_2(\text{SNR}).$$

In terms of  $r_l$ , the cardinality of the ST code  $\mathcal{X}_{\text{ARQ}}$  is given simultaneously by

$$|\mathcal{X}_{\text{ARQ}}| = \text{SNR}^{T r_l}, \quad 1 \leq l \leq L$$

from which it follows that

$$l r_l = j r_j \quad (7)$$

for all  $1 \leq j, l \leq L$  and, in particular,  $r_l = r_1/l$ . The rates  $r_l, R_l$  are to be distinguished from the (average or effective) rate  $R$  and normalized rate  $r$  of the ARQ scheme obtained by taking into consideration the number of rounds of ARQ required and their probability.

We will specifically refer to  $R_1$  and  $r_1$  as the single-round and normalized single-round rates, respectively. The relationship between the rates  $R$  and  $R_1$  plays a key role and will be treated later. The normalized rate  $r$  will also be called the spatial-multiplexing gain of the ARQ scheme.

Let  $P_{e,l}(r_l)$  denote the probability of error of the ST code  $\mathcal{X}_{\text{ARQ},l}$  decoded using decoder  $\mathcal{D}_l$ . In computing  $P_{e,l}(r_l)$ , we consider an error to have occurred only if the decoder  $\mathcal{D}_l$  proceeds to decode, generates an ACK signal and the decoding is in error. Thus, we do not consider the decoder to have made an error if a NACK is generated. The probability of error  $P_e(r)$  of the ARQ code is upper-bounded by

$$P_e(r) \leq \sum_{l=1}^L P_{e,l}(r_l). \quad (8)$$

Although the above is a form of the union bound, we offer the following added explanation. Consider two parallel communication systems. In the first fictitious system, the entire matrix  $\theta H \mathbf{X}_L$  is transmitted corresponding to each message symbol. At the receiver end, the bank of  $L$  decoders  $\mathcal{D}_l$  independently

<sup>1</sup>Notice that the quantity  $2^{lR_l T}$  is clearly not always an integer. For simplicity, we employ the notation  $2^{lR_l T}$  in place of the more accurate notation  $\lceil 2^{lR_l T} \rceil$ , keeping in line with the literature in the area. This convention is used throughout the paper.

examine the message symbol each based on the corresponding fragment  $\mathbf{Y}_l$  of the received signal and generate an ACK and a decoded message symbol or a NACK as appropriate. The second system is the ARQ system under study here. We assumed that both communication systems are faced with the same realization of channel fading coefficients and noise. Clearly, the number of incorrect decodings by the second (ARQ) system cannot exceed the sum of the incorrect decodings by the bank of  $L$  decoders in the fictitious system. Equation (8) now follows.

The sufficiency condition can now be stated.

*Theorem 2:* Let  $\mathcal{X}_{\text{ARQ}}$  be an ST code designed for the MIMO-ARQ channel having normalized single-round rate  $r_1$  and multiplexing gain  $r$ . If the code  $\mathcal{X}_{\text{ARQ}}$ , in conjunction with receiver decoding algorithms  $\{\mathcal{D}_l\}$  satisfies the following.

- 1)  $\Pr(\bar{\mathcal{A}}_1) \doteq \text{SNR}^{-T}$ , where  $T > 0$  for all  $0 \leq r_1 < \min(n_t, n_r)$ .
- 2) The full-length ST code  $\mathcal{X}_{\text{ARQ},L}$  is a D-MG optimal ST code for multiplexing gain  $r_L = r_1/L$ .
- 3)

$$P_{e,l}(r_1) \dot{\leq} P_{e,L}(r_L), \quad 1 \leq l \leq (L-1). \quad (9)$$

Then

$$\lim_{\text{SNR} \rightarrow \infty} r = r_1$$

and the ST code  $\mathcal{X}_{\text{ARQ}}$  achieves the DMD tradeoff for all  $r$ ,  $0 \leq r < \min\{n_t, n_r\}$ .

In words, the theorem asserts that if an ST-ARQ code  $\mathcal{X}_{\text{ARQ}}$  satisfies the requirements:

- 1) that with high probability there will be just a single ARQ round and
- 2) the full-length ST code  $\mathcal{X}_{\text{ARQ},L}$  is optimal with respect to the D-MG tradeoff of the channel,
- 3) the error probability of the  $l$ th decoder  $\mathcal{D}_l$  applied to the task of decoding the ST code  $\mathcal{X}_{\text{ARQ},l}$  is no larger than that incurred by the ML decoder  $\mathcal{D}_L$  applied to the task of decoding the ST code  $\mathcal{X}_{\text{ARQ},L}$

then  $\mathcal{X}_{\text{ARQ}}$  will achieve the DMD tradeoff.

*Proof:* (Theorem 2)

a) *Showing that  $R \rightarrow R_1$  for high SNR:* For  $1 \leq l \leq L$ , let  $p(l) = P_r(\bar{\mathcal{A}}_1, \dots, \bar{\mathcal{A}}_l)$ , be the probability of the transmitter receiving a NACK in each of the first  $l$  ARQ rounds. Note that  $p(L) = 0$ . Then from [1] the rate  $R$  of the ARQ scheme and the single-round rate  $R_1$  are related by

$$R = \frac{R_1}{1 + \sum_{l=1}^{L-1} p(l)}. \quad (10)$$

The probabilities  $p(l)$ ,  $l = 1, 2, \dots, (L-1)$ , can be upper-bounded as follows:

$$\begin{aligned} p(l) &= \Pr(\bar{\mathcal{A}}_1, \dots, \bar{\mathcal{A}}_l) \\ &\leq \Pr(\bar{\mathcal{A}}_1) \\ &\doteq \text{SNR}^{-T}, \quad T > 0, \quad \forall r_1 < \min\{n_t, n_r\}. \end{aligned}$$

The last equality follows from theorem hypotheses. It follows now from (10) that

$$\frac{R_1}{1 + (L-1)\text{SNR}^{-T}} \leq R \leq R_1.$$

As a result, we have

$$\lim_{\text{SNR} \rightarrow \infty} R = R_1. \quad (11)$$

Since  $R = r \log_2(\text{SNR})$  and  $R_1 = r_1 \log_2(\text{SNR})$ , it follows that

$$\lim_{\text{SNR} \rightarrow \infty} r = r_1 \quad (12)$$

as well. Thus, at large SNR, the single-round and average rates  $R_1$  and  $R$ , respectively, as well as the corresponding normalized versions  $r_1, r$ , may be regarded as being equal.

a) *Lower bound on the probability of error:* Let  $P_e(r)$  denote the probability of a message symbol being decoded incorrectly when  $\mathcal{X}_{\text{ARQ}}$  is used in conjunction with some decoding algorithm. (Note that although we can have any decoding algorithm for  $l < L$ ,  $\mathcal{D}_L$  has been fixed to ML.) Then we have

$$P_e(r) \geq P_{e,L}(r_L). \quad (13)$$

b) *Upper bound on the probability of error:* If a ARQ scheme satisfies condition (9), from (8) we have

$$P_e(r) \dot{\leq} P_{e,L}(r_L).$$

Using the fact that full-length ST code is D-MG optimal in conjunction with the above bounds we have

$$P_e(r) \doteq \text{SNR}^{-d^*(r/L)}$$

i.e.,

$$P_e(r) \doteq \text{SNR}^{-d^*(r/L)}$$

from (7) and (12).  $\square$

### III. MINIMUM-DELAY CONSTRUCTION OF DMD OPTIMAL ARQ ST SCHEMES

The constructions presented in this section achieve minimum delay and can be applied whenever  $n_r \geq n_t$ , the channel is a regular fading channel, and whenever either  $L|n_t$  or  $n_t|L$ . All constructions presented in this paper are based on cyclic division algebras (CDA) and we begin with a brief review of ST code construction from CDAs.

#### A. ST Codes Derived From CDAs

The construction of ST codes from division algebras was first proposed by Sethuraman and Sundar Rajan [13].

Division algebras are noncommutative rings with identity element and inverses, i.e., each nonzero element has a multiplicative inverse. CDAs have a particularly simple structure and a general technique for the construction of a CDA can be found in [18], [14, Proposition 11], or [19, Theorem 1]. Let  $\mathbb{F}, \mathbb{L}$  be number fields, with  $\mathbb{L}$  a finite, cyclic Galois extension of  $\mathbb{F}$  of degree  $n$ , see Fig. 2. Let  $\sigma$  denote a 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  $\gamma \in \mathbb{F}^*$ , i.e., some element  $\gamma$  having the property that the smallest positive integer  $t$  for which  $\gamma^t$  is

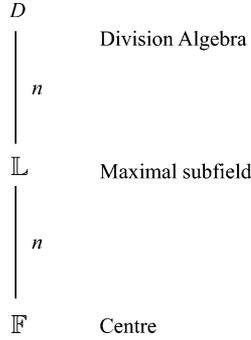


Fig. 2. Cyclic division algebra.

the relative norm<sup>2</sup>  $N_{\mathbb{L}/\mathbb{F}}(u)$  of some element  $u$  in  $\mathbb{L}^*$ , is  $n$ . Then the set of all elements of the form

$$\sum_{i=0}^{n-1} z^i \ell_i \quad (15)$$

forms a CDA  $D(\mathbb{L}/\mathbb{F}, \sigma, \gamma)$  with center  $\mathbb{F}$  and maximal subfield  $\mathbb{L}$ .

It can be verified that  $D$  is a (right) vector space over the maximal subfield  $\mathbb{L}$ . The parameter  $n$  is called the *index* of the CDA. An ST code  $\mathcal{X}_{n \times n}$  can be associated to  $D$  by selecting the set of matrices corresponding to a finite subset of  $D$ . The matrix associated to an element  $d \in D$  corresponds to the left multiplication by the element  $d$  in the division algebra. Let  $\lambda_d$  denote this operation,  $\lambda_d : D \rightarrow D$ , defined by

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

It can be verified that  $\lambda_d$  is a  $\mathbb{L}$ -linear transformation of  $D$ . From (15), a natural choice of basis for the right-vector space  $D$  over  $\mathbb{L}$  is  $\{1, z, z^2, \dots, z^{n-1}\}$ . A typical element in the division algebra  $D$  is  $d = \ell_0 + z\ell_1 + \dots + z^{n-1}\ell_{n-1}$ , where the  $\ell_i \in \mathbb{L}$ . The matrix representation of the  $\mathbb{L}$ -linear transformation  $\lambda_d$  under this basis can be shown to be [14], [19]

$$\begin{bmatrix} \ell_0 & \gamma\sigma(\ell_{n-1}) & \gamma\sigma^2(\ell_{n-2}) & \dots & \gamma\sigma^{n-1}(\ell_1) \\ \ell_1 & \sigma(\ell_0) & \gamma\sigma^2(\ell_{n-1}) & \dots & \gamma\sigma^{n-1}(\ell_2) \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \ell_{n-1} & \sigma(\ell_{n-2}) & \sigma^2(\ell_{n-3}) & \dots & \sigma^{n-1}(\ell_0) \end{bmatrix}. \quad (16)$$

Our ST codes are derived from matrices of the form in (16), where the elements  $\ell_i$  are restricted to be of the form

$$\ell_i = \sum_{k=1}^n e_{i,k} \beta_k, \quad e_{i,k} \in \mathcal{A}_{\text{QAM}} \quad (17)$$

with

$$\mathcal{A}_{\text{QAM}} = \{a + ib \mid -M + 1 \leq a, b \leq M - 1, a, b \text{ odd}\}$$

<sup>2</sup>The ‘‘relative norm’’  $N_{\mathbb{L}/\mathbb{F}}(u)$  of an element  $u \in \mathbb{L}$  is given by

$$N_{\mathbb{L}/\mathbb{F}}(u) = \prod_{i=0}^{n-1} \sigma^i(u) \quad (14)$$

and it can be verified that  $N_{\mathbb{L}/\mathbb{F}}(u) \in \mathbb{F}$ .

and where  $\beta_k, k = 1, 2, \dots, n$  is an integral basis (i.e., a basis as a module) for  $\mathcal{O}_{\mathbb{L}}/\mathcal{O}_{\mathbb{F}}$ . We will refer to the underlying  $M^2$ -QAM-alphabet  $\mathcal{A}_{\text{QAM}}$  as the *base alphabet* of the ST code construction. Note that  $|\mathcal{A}_{\text{QAM}}| = M^2$ . Let  $\mathcal{X}_{n \times n}$  denote the collection of all matrices of the form given in (16), (17). Clearly, this code has cardinality  $[M^2]^{n^2}$ . If it is desired to communicate at a rate  $r \log_2(\text{SNR})$  bits per channel use, then one must choose the size  $M^2$  of the underlying  $\mathcal{A}_{\text{QAM}}$  alphabet accordingly, i.e., set

$$M^2 = \text{SNR}^{r/n}. \quad (18)$$

Note that  $|e_{i,k}|^2 \leq 2(M-1)^2 \doteq M^2 \doteq \text{SNR}^{r/n}$ . Thus

$$|\ell_i|^2 = \left| \sum_{k=1}^n e_{i,k} \beta_k \right|^2 \leq M^2 \quad (19)$$

with the latter inequality following from the fact that the basis elements  $\{\beta_i\}$  have magnitude that is independent of the SNR. Let  $\Delta X$  denote the difference of any two distinct codeword matrices from  $\mathcal{X}_{n \times n}$ . A key property of this construction, established in [10] using properties of the underlying CDA, is that

$$\min_{\Delta X} \det(\theta^2 \Delta X \Delta X^\dagger) \doteq \text{SNR}^{n_t - r}. \quad (20)$$

From Theorem 1, it follows that the square ( $T = n_t$ ) ST codes constructed from CDA as described above, achieve the D-MG tradeoff for all fading distributions.

### B. Decoding Algorithm

Our minimum-delay ST constructions will be shown to be DMD optimal when used in conjunction with a bounded distance decoder, the operation of which we will present first. Following this, we will present the minimum delay construction for the cases of  $L|n_t$  and  $n_t|L$ , and show that these are DMD optimal for  $n_r \geq n_t$  according to Theorem 2.

Without loss of generality, we will identify the integer set  $\{w : 1 \leq w \leq |\mathcal{X}_{\text{ARQ}}|\}$  with the collection of  $|\mathcal{X}_{\text{ARQ}}|$  message symbols. As discussed above, the ARQ scheme employs a total of  $L$  decoders  $\{\mathcal{D}_l \mid 1 \leq l \leq L\}$  corresponding to the  $L$  ARQ rounds. In our analysis, we will assume that the decoders employed in the first  $L-1$  rounds are bounded-distance decoders, operating as described below (see also Fig. 3).

The decoder  $\mathcal{D}_L$  employed at the end of the  $L$ th round, as stated earlier, is assumed to be an ML decoder. It is convenient to regard the decoders  $\mathcal{D}_l$  as mappings.

Case (i):  $l = 1, 2, \dots, (L-1)$

- $\mathcal{D}_l(\mathbf{Y}_l) = \hat{w}$  for some message symbol  $\hat{w}$ ,  $1 \leq \hat{w} \leq |\mathcal{X}_{\text{ARQ}}|$  if the codeword  $\theta \hat{\mathbf{X}}_l$  in  $\mathcal{X}_{\text{ARQ},l}$  corresponding to message symbol  $\hat{w}$  is the unique codeword such that

$$\|\mathbf{Y}_l - \theta H \hat{\mathbf{X}}_l\|_{\mathbb{F}}^2 \leq n_r T l (1 + \delta)$$

where  $\delta = \beta \log(\text{SNR})$  for some  $\beta > 0$ . Here  $\beta$  is a suitably chosen parameter of the bounded distance decoder. The receiver sends out an ACK in this case.

- $\mathcal{D}_l(\mathbf{Y}_l) = 0$  in any other case indicating that the decoder is unable to decode with some predetermined level

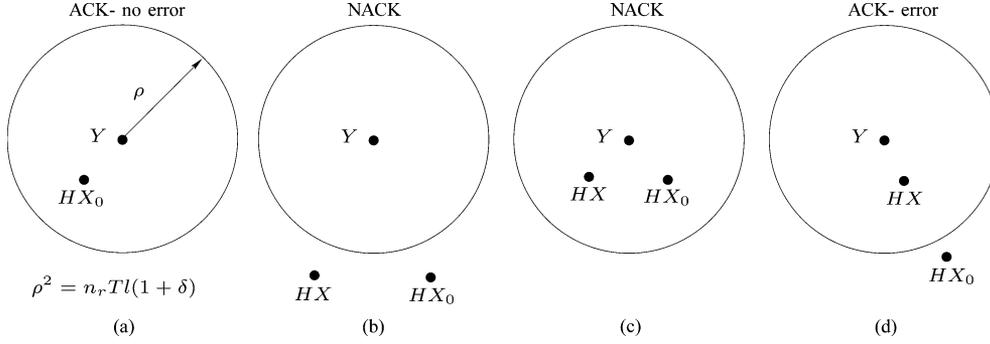


Fig. 3. Bounded-distance decoding in intermediate rounds and associated error events: (a) Transmitted code matrix is the only code-matrix within the sphere, ACK is generated, no error. (b) No code matrix is inside the sphere, NACK is generated. (c) Two (or more) code matrices within the sphere, NACK is generated. (d) A single code matrix which is not the transmitted matrix is within the sphere, ACK is generated, error.

of confidence to a codeword. The receiver sends out a NACK in this case.

Case (ii):  $l = L$

- $\mathcal{D}_L$  is the mapping corresponding to ML decoding, i.e., to choosing the message  $\hat{w}$  maximizing  $\Pr(\mathbf{Y}_L | \hat{w})$ ,  $\hat{w} = 1, \dots, |\mathcal{X}_{\text{ARQ}}|$ . Since ML decoding will always result in a decoding decision, implicitly an ACK is always generated following the conclusion of the  $L$ th ARQ round causing the transmitter to move on to transmitting the next message symbol.

The following lemma is key to our code construction.

*Lemma 3:* Let  $\mathcal{X}_{\text{ARQ}}$  be an ST code for an  $L$ -round multiple-input multiple-output (MIMO)-ARQ channel and let  $\mathcal{X}_{\text{ARQ},1}$  denote the single-round ST code as defined in Section I. Let  $\theta\Delta X_1$  denote the difference of any two distinct matrices from  $\mathcal{X}_{\text{ARQ},1}$  and set

$$\min_{\Delta X_1} \theta^2 \|\Delta X_1\|_F^2 \doteq \text{SNR}^T.$$

Then, for the case when  $n_r \geq n_t$ , the probability of the transmitter receiving a NACK at the conclusion of the first round from a receiver that employs the bounded-distance decoder  $\mathcal{D}_1$  described previously in Section III-B, can be upper-bounded by

$$\Pr(\bar{\mathcal{A}}_1) \dot{\leq} \text{SNR}^{-T}.$$

*Proof:* See Appendix A for a proof. It is in the proof of this lemma that use is made of the fact that communication takes place over a regular fading channel.  $\square$

The theorem that follows is a restatement of Theorem 2 for the case when the decoder  $\mathcal{D}_1$  is a bounded-distance decoder.

*Theorem 4:* Let  $n_r \geq n_t$  and  $\mathcal{X}_{\text{ARQ}}$ ,  $\mathcal{X}_{\text{ARQ},1}$ ,  $\theta\Delta X_1$ ,  $T$  be as in Lemma 3. Then the ST code  $\mathcal{X}_{\text{ARQ}}$  achieves the optimal DMD tradeoff for all  $r$ ,  $0 \leq r < n_t$ , if

1) *Frobenius-Norm Criterion:*

$$\min_{\Delta X_1} \theta^2 \|\Delta X_1\|_F^2 \doteq \text{SNR}^T$$

with  $T > 0$  for all  $0 \leq r_1 < n_t$ , and

- 2) *D-MG Optimality Criterion:* The full-length ST code  $\mathcal{X}_{\text{ARQ},L}$  is D-MG optimal, and
- 3) *Error-Probability Criterion*

$$P_{e,l}(r_l) \dot{\leq} P_{e,L}(r_L), \quad 1 \leq l \leq (L-1).$$

*Proof:* Immediate from Theorem 2 and Lemma 3.  $\square$

Constructions for DMD-optimal ARQ ST codes are presented for cases (i)  $L|n_t$  and (ii)  $n_t|L$  in Sections III-C and III-D, respectively.

### C. Construction for the Case $L|n_t$

Let the block length  $T$  be given by  $T = n_t/L$ . The construction of DMD-optimal ARQ codes in this case is derived from a square  $(LT \times LT)$  ST code  $\mathcal{X}_{(LT \times LT)}$  obtained via the construction in (16) with  $n = LT$ . Every codeword  $X \in \mathcal{X}_{(LT \times LT)}$  is of the form

$$X = \begin{bmatrix} \ell_0 & \gamma\sigma(\ell_{n-1}) & \cdots & \gamma\sigma^{n-1}(\ell_1) \\ \ell_1 & \sigma(\ell_0) & \cdots & \gamma\sigma^{n-1}(\ell_2) \\ \ell_2 & \sigma(\ell_1) & \cdots & \gamma\sigma^{n-1}(\ell_3) \\ \vdots & \vdots & \ddots & \vdots \\ \ell_{n-1} & \sigma(\ell_{n-2}) & \cdots & \sigma^{n-1}(\ell_0) \end{bmatrix}_{n \times n}.$$

Now if  $R = r \log_2(\text{SNR})$  is the desired rate of the ARQ scheme, we set

$$|\mathcal{X}_{(LT \times LT)}| = \text{SNR}^{Tr}. \quad (21)$$

Each round of the ARQ transmission corresponds to transmitting  $T$  successive columns from the matrix  $X$ , i.e., during the  $l$ th round, we transmit  $\theta X_l = \theta[\underline{c}_{(l-1)T+1} \cdots \underline{c}_{lT}]$  for  $l = 1, 2, \dots, L$ , where  $\underline{c}_i$  denotes the  $i$ th column of  $X$ .

*Example 1:* Consider the case when  $n_t = 4, L = 2$ . This leads to the choice  $T = n_t/L = 2$ . Each codeword matrix  $X \in \mathcal{X}_{(4 \times 4)}$  is of the form

$$X = \begin{bmatrix} \ell_0 & \gamma\sigma(\ell_3) & \gamma\sigma^2(\ell_2) & \gamma\sigma^3(\ell_1) \\ \ell_1 & \sigma(\ell_0) & \gamma\sigma^2(\ell_3) & \gamma\sigma^3(\ell_2) \\ \ell_2 & \sigma(\ell_1) & \sigma^2(\ell_0) & \gamma\sigma^3(\ell_3) \\ \ell_3 & \sigma(\ell_2) & \sigma^2(\ell_1) & \sigma^3(\ell_0) \end{bmatrix}_{4 \times 4}.$$

In the above construction, we choose  $\mathbb{L}$  to be the cyclotomic field obtained by adjoining the 16th root of unity to the rationals  $\mathbb{Q}$ , i.e., we set  $\mathbb{L} = \mathbb{Q}(\omega_{16})$ ,  $\omega = e^{i2\pi/16}$ ,  $i = \sqrt{-1}$ . It can be shown that  $\gamma = 2 + i$  is a valid non-norm element. One choice for the generator  $\sigma$  of the Galois group  $\text{Gal}(\mathbb{L}/\mathbb{Q}(i))$  is the automorphism  $\sigma : \omega_{16} \mapsto \omega_{16}^5$ . The elements  $\ell_i$  take on values from the ring of integers  $\mathbb{Z}[\omega_{16}]$  in accordance with (17). The matrices  $X_1, X_2$  associated with transmissions during the two rounds of ARQ transmission are derived from  $X$  above by

selecting  $X_1, X_2$  to be comprised of the first two and last two columns of  $X$  respectively, i.e.,

$$X_1 = \begin{bmatrix} \ell_0 & \gamma\sigma(\ell_3) \\ \ell_1 & \sigma(\ell_0) \\ \ell_2 & \sigma(\ell_1) \\ \ell_3 & \sigma(\ell_2) \end{bmatrix}, \quad X_2 = \begin{bmatrix} \gamma\sigma^2(\ell_2) & \gamma\sigma^3(\ell_1) \\ \gamma\sigma^2(\ell_3) & \gamma\sigma^3(\ell_2) \\ \sigma^2(\ell_0) & \gamma\sigma^3(\ell_3) \\ \sigma^2(\ell_1) & \sigma^3(\ell_0) \end{bmatrix}.$$

**Theorem 5:** The ST code  $\mathcal{X}_{\text{ARQ}}$  constructed above for the case  $L|n_t$  achieves the MIMO-ARQ DMD tradeoff of the long-term-static ARQ channel for  $n_r \geq n_t$  under the power constraint (3), i.e.,

$$d_{\mathcal{X}_{\text{ARQ}}}(r, L) = d^* \left( \frac{r}{L} \right), \quad 0 \leq r < n_t$$

for block length  $T = n_t/L$ .

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

It is shown in Section III-E that the above construction possesses the smallest possible value of delay parameter  $T$ .

#### D. Construction for the Case $n_t|L$

The constructions in this case, are derived from the constructions of D-MG optimal rectangular ST codes presented in [10]. Two constructions of D-MG optimal ST codes were presented in [10]: row deletion and horizontal stacking. Note that the idea of using horizontal stacking to construct rectangular ST codes was first presented in [4].

ARQ constructions derived from these rectangular codes permit setting  $T = 1$  which is clearly the minimum possible.<sup>3</sup>

1) *Row-Deletion ARQ Construction:* As indicated above, we set  $T = 1$ . Let  $L = n_t k$ , for some integer  $k$ . The row-deletion ST code for the ARQ channel is derived starting from the corresponding row-deletion D-MG-optimal  $(n_t \times L)$  rectangular construction [10], that is obtained as follows. An  $n_t \times L$  row-deleted rectangular ST code  $\mathcal{X}_{n_t \times L}$  is obtained in [10] by removing the last  $L - n_t$  rows from an  $L \times L$  square ST code of the form given in (16). It is shown in [10] that  $\mathcal{X}_{n_t \times L}$  also satisfies Theorem 1 and is hence D-MG optimal. Suppose that  $\mathcal{X}_{n_t \times L}$  has been constructed using a division algebra  $D$  having center  $\mathbb{F}$  and maximal subfield  $\mathbb{L}$  as shown in Fig. 4(a). Each codeword  $X \in \mathcal{X}_{n_t \times L}$  is an  $n_t \times L$  matrix of the form

$$X = \begin{bmatrix} \ell_0 & \gamma\sigma(\ell_{L-1}) & \cdots & \gamma\sigma^{L-1}(\ell_1) \\ \ell_1 & \sigma(\ell_0) & \cdots & \gamma\sigma^{L-1}(\ell_2) \\ \ell_2 & \sigma(\ell_1) & \cdots & \gamma\sigma^{L-1}(\ell_3) \\ \vdots & \vdots & \ddots & \vdots \\ \ell_{n_t-2} & \sigma(\ell_{n_t-3}) & \cdots & \gamma\sigma^{L-1}(\ell_{n_t-1}) \\ \ell_{n_t-1} & \sigma(\ell_{n_t-2}) & \cdots & \gamma\sigma^{L-1}(\ell_{n_t}) \end{bmatrix}. \quad (22)$$

We identify a set of  $k-1$  columns of the matrix  $X$ , which along with the first column will contain all the independent variables  $\{\ell_i\}_{i=0}^{L-1}$ . Let  $\mathbb{L}'$  be a degree- $k$  extension of the maximal subfield  $\mathbb{L}$  that has an integral basis  $\{\alpha_i\}_{i=1}^k$  over  $\mathbb{L}$ , as shown in Fig. 4(b). Such an extension can always be found. We set  $\alpha_1 = 1$  without loss of generality. Thus, the  $\{\alpha_i\}_{i=1}^k$  will lie in the in-

<sup>3</sup>The construction can be extended to the case of arbitrary  $T$ , but clearly setting  $T = 1$  results in minimum delay.

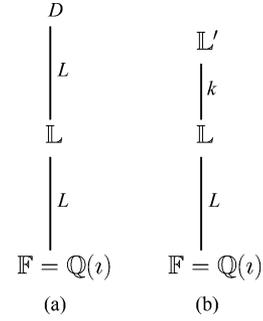


Fig. 4. Algebraic tower for the row-deleted ARQ construction.

tegral closure  $\mathcal{O}_{\mathbb{L}'}$  of  $\mathcal{O}_{\mathbb{L}}$  in  $\mathbb{L}'$ . Consider a column vector  $\underline{c}$  obtained by taking a linear combination of these  $k$  columns with coefficients  $\{\alpha_i\}_{i=1}^k$ . Since the transmitted message symbol can be recovered from the elements  $\{\ell_i\}_{i=0}^{L-1}$ , column vector  $\underline{c}$  has the same information content as does the entire ST code matrix  $X$ . Next, replace the first column of  $X$  with column  $\underline{c}$ . The above procedure of combining  $k-1$  columns into the first column is an elementary column operation that does not change either the determinant or the SNR exponent of the Frobenius norm (energy) of the  $(n_t \times L)$  ST matrix  $X$ . As a consequence, it follows from the results in [10] that this modified  $(n_t \times L)$  ST code is also optimal with respect to the D-MG tradeoff. We now define the  $n_t \times L$  ARQ ST scheme  $\mathcal{X}_{\text{ARQ}}$  as one in which the  $L$  ARQ rounds involve transmission of the  $L$  columns of the modified  $(n_t \times L)$  ST code matrix in turn beginning with the first. The following example illustrates this procedure.

*Example 2:* Consider the case when  $n_t = 2$ ,  $L = 4$ ,  $T = 1$ ,  $k = L/n_t = 2$ . We start by constructing a  $(2 \times 4)$  rectangular row-deleted D-MG optimal ST code  $\mathcal{X}_{2 \times 4}$  from CDA as in [10]. Each codeword  $X \in \mathcal{X}_{2 \times 4}$  is of the form

$$X = \begin{bmatrix} \ell_0 & \gamma\sigma(\ell_3) & \gamma\sigma^2(\ell_2) & \gamma\sigma^3(\ell_1) \\ \ell_1 & \sigma(\ell_0) & \gamma\sigma^2(\ell_3) & \gamma\sigma^3(\ell_2) \end{bmatrix}_{2 \times 4}.$$

The above construction is derived by deleting the last two rows of a  $(4 \times 4)$  square CDA-based construction such as the one provided in Example 1. As an example, we can take  $\mathbb{L} = \mathbb{Q}(\omega_{16})$ , non-norm element  $\gamma = 2 + i$  and generator  $\sigma$  of the Galois group  $\text{Gal}(\mathbb{L}/\mathbb{Q}(i))$  to be the automorphism  $\sigma : \omega_{16} \mapsto \omega_{16}^5$ . The elements  $\ell_i$  take on values from the ring of integers  $\mathbb{Z}[\omega_{16}]$  as in (17). To construct the ARQ ST code, we first identify the first and the third columns as the  $k = 2$  columns that contain all independent (i.e., message-bearing) variables. Let  $\{\alpha_1, \alpha_2\}$  constitute an integral basis of the degree-2 extension  $\mathbb{L}' = \mathbb{Q}(\omega_{32})$  of  $\mathbb{L}$ . Perform the elementary column operation corresponding to replacing the first column of the matrix  $X$  by the sum of  $\alpha_1$  times the first column and  $\alpha_2$  times the third column to obtain the matrix

$$\begin{bmatrix} \ell_0\alpha_1 + \alpha_2\gamma\sigma^2(\ell_2) & \gamma\sigma(\ell_3) & \gamma\sigma^2(\ell_2) & \gamma\sigma^3(\ell_1) \\ \ell_1\alpha_1 + \alpha_2\gamma\sigma^2(\ell_3) & \sigma(\ell_0) & \gamma\sigma^2(\ell_3) & \gamma\sigma^{T-1}(\ell_2) \end{bmatrix}_{2 \times 4}.$$

Our  $n_t \times L$  ST scheme for the MIMO-ARQ channel then simply involves transmitting the  $i$ th column,  $i = 1, 2, 3, 4$  of the above matrix during the  $i$ th round of transmission.

*Theorem 6:* The row-deleted ARQ ST code  $\mathcal{X}_{\text{ARQ}}$  constructed above for the case when  $n_t|L$  achieves the MIMO-ARQ DMD tradeoff of the long-term static ARQ channel for  $n_r \geq n_t$  under the power constraint (3), i.e.,

$$d_{\mathcal{X}_{\text{ARQ}}}(r, L) = d^* \left( \frac{r}{L} \right), \quad 0 \leq r < n_t$$

for block length  $T = 1$ .

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

2) *Horizontal-Stacking ARQ Construction:* The horizontally stacked-ARQ ST construction is very similar to that of the row-deleted ARQ construction, except that the starting point is a horizontally stacked rectangular construction [4], [10]. An  $n_t \times L$  horizontally stacked rectangular construction is obtained in [10] by horizontally stacking  $k$  number of  $n_t \times n_t$  square CDA ST codes. This construction is also D-MG optimal [10]. To construct a horizontally stacked-ARQ ST code, we start from a horizontally stacked rectangular construction and use elementary column operations similar to those used in Section III-D.1 to ensure that all independent variables are accommodated in the first round of transmission.

*Example 3:* Consider the case  $n_t = 2$ ,  $L = 4$ ,  $T = 1$ . The horizontally stacked rectangular ( $2 \times 4$ ) construction takes on the form

$$\begin{bmatrix} \ell_0 & \gamma\sigma(\ell_1) & \ell_2 & \gamma\sigma(\ell_3) \\ \ell_1 & \sigma(\ell_0) & \ell_3 & \sigma(\ell_2) \end{bmatrix}.$$

In this example, the  $\ell_i$  are drawn from the ring of integers  $\mathbb{Z}[\omega_8]$  of the cyclotomic extension  $\mathbb{L} = \mathbb{Q}(\omega_8)$  according to (17). The non-norm element  $\gamma = 2 + \iota$  is chosen and the Galois-group generator  $\sigma$  is given by  $\sigma : \omega_8 \mapsto \omega_8^5$ . Let  $\{\alpha_1, \alpha_2\}$  be an integral basis for the field  $\mathbb{L}' = \mathbb{Q}(\omega_{16})$  over  $\mathbb{L} = \mathbb{Q}(\omega_8)$ . We arrive at the following construction through performing an elementary column operation on the above horizontally stacked rectangular construction:

$$\begin{bmatrix} \alpha_1\ell_0 + \alpha_2\ell_2 & \gamma\sigma(\ell_1) & \ell_2 & \gamma\sigma(\ell_3) \\ \alpha_1\ell_1 + \alpha_2\ell_3 & \sigma(\ell_0) & \ell_3 & \sigma(\ell_2) \end{bmatrix}.$$

The horizontally stacked-ARQ construction corresponds to transmitting the  $i$ th column of the above matrix during the  $i$ th round,  $i = 1, 2, 3, 4$ .

*Theorem 7:* The horizontally stacked-ARQ ST code  $\mathcal{X}_{\text{ARQ}}$  constructed above for the case when  $n_t|L$  achieves the MIMO-ARQ DMD tradeoff of the long-term static ARQ channel for  $n_r \geq n_t$  under the power constraint (3), i.e.,

$$d_{\mathcal{X}_{\text{ARQ}}}(r, L) = d^* \left( \frac{r}{L} \right), \quad 0 \leq r < n_t$$

with block length  $T = L/n_t$ .

*Proof:* Similar to the proof of Theorem 6 provided in Appendix C.  $\square$

#### E. Delay Optimality of the Constructions for $L|n_t$ and $n_t|L$

The constructions of ST codes presented in this section are of minimal delay, where by minimal, we mean that the DMD performance cannot be improved by passing to a larger value of

delay parameter  $T$  and in addition, any smaller value of  $T$  will result in DMD-performance degradation.

*Theorem 8:* The constructions presented for the case of  $L|n_t$  and  $n_t|L$  are delay-optimal and achieve the minimum possible block length of

$$T = \left\lceil \frac{n_t}{L} \right\rceil.$$

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

*Remark 1:* In [1], the authors construct finite block-length DMD optimal random Gaussian and IR-LAST codes, provided that they respectively satisfy

$$T \geq \left\lceil \frac{n_t + n_r - 1}{L} \right\rceil$$

and

$$T \geq \lceil n_t + n_r - 1 \rceil.$$

The codes in this section, for the cases  $n_t|L$  or  $L|n_t$ , achieve optimality with a significantly smaller value of  $T$ .

## IV. A GENERAL, APPROXIMATELY UNIVERSAL CONSTRUCTION

The construction presented in this section, which while not guaranteeing minimum delay, has the advantage of being widely applicable:

- the construction is approximately universal [8] and hence generates ST codes that are DMD optimal for any statistical description of the fading channel;
- can be applied both for  $n_r < n_t$  or  $n_r \geq n_t$ ;
- there are no restrictions on the number  $L$  of ARQ rounds.

The construction of the code and the manner in which it is used are quite different from that of the constructions in the previous section. While still based on the theory of cyclic division algebras, the construction in this section is an adaptation of the construction of approximately universal ST codes for the block-fading channel, see [15], [23], [24]. In the construction, the block length  $T = n_t$  and we will use the integer  $n$  to denote their common value, i.e.,  $n = T = n_t$  in this section.

### A. Constructing the Appropriate Cyclic Division Algebra

Let  $m \geq L$  be the smallest integer such that the greater common divisor (gcd)  $(m, n)$  of  $m, n$  equals 1. Let  $\mathbb{K}, \mathbb{M}$  be cyclic Galois extensions of  $\mathbb{Q}(\iota)$  of degrees  $m, n$  whose Galois groups are generated, respectively, by the automorphisms  $\phi_1, \sigma_1$ , i.e.,

$$\begin{aligned} \text{Gal}(\mathbb{K}/\mathbb{Q}(\iota)) &= \langle \phi_1 \rangle \\ \text{Gal}(\mathbb{M}/\mathbb{Q}(\iota)) &= \langle \sigma_1 \rangle. \end{aligned}$$

Let  $\mathbb{L}$  be the composite of  $\mathbb{K}, \mathbb{M}$ , see Fig. 5. Then it is known that  $\mathbb{L}/\mathbb{Q}(\iota)$  is cyclic and that further

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

Thus, every element of  $\text{Gal}(\mathbb{L}/\mathbb{Q}(\iota))$  can be associated with a pair  $(\phi_1^i, \sigma_1^j)$  belonging to  $\text{Gal}(\mathbb{K}/\mathbb{Q}(\iota)) \times \text{Gal}(\mathbb{M}/\mathbb{Q}(\iota))$ . Let  $\phi, \sigma$  be the automorphisms associated to the pairs  $(\phi_1, \text{id})$ ,  $(\text{id}, \sigma_1)$ , respectively, where  $\text{id}$  denotes the corresponding

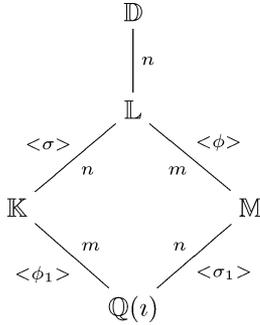


Fig. 5. Construction of the underlying cyclic-division algebra.

identity automorphism. Then  $\phi, \sigma$  are the generators of the Galois groups  $\text{Gal}(\mathbb{L}/\mathbb{M})$ ,  $\text{Gal}(\mathbb{L}/\mathbb{K})$ , respectively.

Let  $\gamma \in \mathbb{K}$  be a non-norm element of the extension  $\mathbb{L}/\mathbb{K}$ . Let  $z$  be an indeterminate satisfying  $z^n = \gamma$ . Consider the  $n$ -dimensional vector space

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

We define multiplication on  $D$  by setting  $\ell_i z = z\sigma(\ell_i)$  and, as before, this turns  $D$  into a CDA whose center is  $\mathbb{K}$  and having  $\mathbb{L}$  as a maximal subfield. Given a matrix  $X$  with components  $X_{i,j} \in \mathbb{L}$ , we define  $\phi(X)$  to be the matrix over  $\mathbb{L}$  whose  $(i, j)^{\text{th}}$  component is given by  $[\phi(X)]_{i,j} = \phi([X]_{i,j})$ . Note that in this case

$$\begin{aligned} \prod_{i=0}^{m-1} \det(\phi^i(X)) &= \prod_{i=0}^{m-1} \phi^i(\det(X)) \\ &= \prod_{i=0}^{m-1} \phi_1^i(\det(X)) \in \mathbb{Q}(i) \end{aligned}$$

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

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

so that

$$\left| \prod_{i=0}^{m-1} \det(\phi^i(X)) \right|^2 \geq 1. \quad (23)$$

### B. ST Code Construction on the CDA

Let  $\mathcal{X}$  be the  $(n \times n)$  ST code comprised of the matrix representations of the elements  $\sum_{i=0}^{n-1} z^i \ell_i$ , where  $\ell_i$  are restricted to be of the form

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

where  $\{\gamma_1, \dots, \gamma_{nm}\}$  are a basis for  $\mathbb{L}/\mathbb{Q}(i)$ . Note that as a result, we have ensured that  $\ell_i \in \mathcal{O}_{\mathbb{L}}$ .

For  $1 \leq l \leq m$ , let  $\mathcal{X}_{\text{ARQ},l}$  be the  $(n \times ln)$  ST code comprised of code matrices having the block form

$$\mathcal{X}_{\text{ARQ},l} = \{\theta[X \phi(X) \cdots \phi^{l-1}(X)] \mid X \in \mathcal{X}\}$$

where, as before,  $\theta$  accounts for SNR normalization. Thus, in reference to our previous notation, we have  $X_l = \phi^{l-1}(X)$ . Although defined for  $1 \leq l \leq m$ , the ARQ scheme will only make use of the codes  $\mathcal{X}_{\text{ARQ},l}$ ,  $1 \leq l \leq L$ . The extended-index notation will, however, prove useful in the proofs. The signal received at the end of the  $l$ th ARQ round, is given by

$$[Y_1 Y_2 \cdots Y_l] = \theta H[X \phi(X) \cdots \phi^{l-1}(X)] + [W_1 W_2 \cdots W_l]. \quad (24)$$

From information-rate considerations, it follows that

$$(M^2)^{mn^2} = \text{SNR}^{nr_1}$$

so that

$$(M^2) = \text{SNR}^{r_1/mn}. \quad (25)$$

It follows that

$$\theta^2 = \text{SNR}^{1-r_1/mn}. \quad (26)$$

### C. Algorithm for Generating Acknowledgments

The generation of an ACK during the  $l$ th ARQ round is based only on whether or not the channel matrix  $H$  is in outage with respect to communication at multiplexing gain  $r_1/l$ . More precisely, following transmission of the  $l$ th block

$$\theta[X \phi(X) \cdots \phi^{l-1}(X)], \quad 1 \leq l < L$$

an ACK is generated by the receiver if and only if the receiver is not in outage associated with a normalized transmission rate of  $r_1/l$ , i.e., if and only if the channel matrix  $H$  is such that

$$\log \det(I + \text{SNR} H^\dagger H) > \frac{r_1}{l} \log(\text{SNR}).$$

*Theorem 9:* The ST code  $\mathcal{X}_{\text{ARQ}}$  constructed above achieves the MIMO-ARQ DMD tradeoff of the long-term static-ARQ channel for any values of the parameters  $n_r, n_t, L$  under the power constraint (3), i.e.,

$$d_{\mathcal{X}_{\text{ARQ}}}(r, L) = d^* \left( \frac{r}{L} \right), \quad 0 \leq r < n_t$$

for block length  $T = n_t$ .

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

## V. SIMULATION RESULTS

In Fig. 6, a comparison of two ARQ ST schemes is presented. The first is the minimum-delay, CDA-based ARQ ST scheme presented in this paper, with  $n_t = n_r = L = 2$ ,  $T = 1$  where  $\mathcal{X}_{\text{ARQ}}$  is given by the  $2 \times 2$  Golden code [6]. The second scheme is the IR-LAST scheme [1] with  $n_t = n_r = L = 2$ ,  $T = 3$ . In order to ensure fair comparison, we choose the radius of the bounded distance decoder so that the effective rates of the CDA-based scheme are no less than those of the IR-LAST code reported in [1]. The effective rates shown in the plot are always those of the CDA-based ARQ scheme at each value of SNR. Also shown for comparison are plots of the coherent outage ( $n_t = 2$ , 4 bits/channel use) and the performance

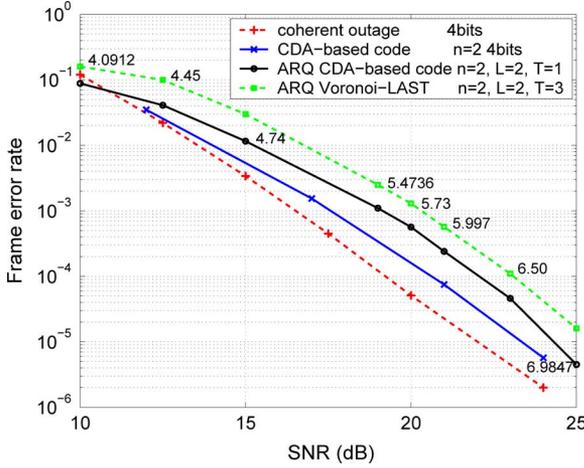


Fig. 6. Average probability of error of CDA based ARQ ST code:  $n_t = n_r = L = 2$ ,  $T = 1$ .

of the full-length CDA-based Golden code. As expected, the ARQ ST-code approaches a rate of 8 bits/channel use at high SNR corresponding to twice the data rate of the full-length code while maintaining a comparable error probability.

#### APPENDIX A PROOF OF LEMMA 3

Let us assume without loss of generality, that  $\theta X_1$  corresponding to message symbol  $w$  is transmitted in the first round. Let  $Y_1$  denote the corresponding received matrix. We define the following events.

- $E_w$ : event that  $\theta H X_1$  is included in a sphere of squared radius  $n_r T(1 + \delta)$  centered around the received matrix  $Y_1$ . We partition this event into the following two sub-events:
  - $E_{1,w}$ : Sub-event that the squared Euclidean distance

$$d_E^2(\theta H X_1, \theta H X'_1)$$

between the matrix  $\theta H X_1$  and its closest neighbor  $\theta H X'_1$  is greater than  $4n_r T(1 + \delta)$ . In this case, the bounded distance decoder  $\mathcal{D}_1$  will decode to the correct message  $w$  and will send an ACK.

- $E_{2,w}$ : Sub-event corresponding to the complement of event  $E_{1,w}$  in  $E_w$ . In this case, the receiver may send either a NACK or an ACK.
- $\bar{E}_w$ : event that  $\theta H X_1$ , corresponding to message  $w_1$ , is not included in a sphere of squared radius  $n_r T(1 + \delta)$  centered around the received matrix  $Y_1$ .

For a given channel realization  $H$ , the probability of a NACK being received at the transmitter at the end of the first round is given by

$$\begin{aligned} \Pr(\bar{\mathcal{A}}_1|H) &= \Pr(E_{2,w} \cap \bar{\mathcal{A}}_1|H) + \Pr(\bar{E}_w \cap \bar{\mathcal{A}}_1|H) \\ &\leq \Pr(E_{2,w}|H) + \Pr(\bar{E}_w|H) \\ &= \Pr(\theta^2 \|H \Delta X_1\|_{F,\min}^2 \leq 4n_r T(1 + \delta)) \\ &\quad + \Pr(\|W_1\|_F^2 \geq n_r T(1 + \delta)) \end{aligned} \quad (27)$$

where  $\theta^2 \|H \Delta X_1\|_{F,\min}^2$  is the minimum squared Euclidean distance between any two distinct codewords in  $\mathcal{X}_{\text{ARQ},1}$  after multiplication by the channel matrix  $H$ . Let  $\lambda_{\min}$  be the minimum

eigenvalue of the matrix  $H^\dagger H$ . Since  $n_r \geq n_t$ , we can write the first term in the preceding expression as

$$\begin{aligned} P_1(H) &:= \Pr(\theta^2 \|H \Delta X_1\|_{F,\min}^2 \leq 4n_r T(1 + \delta)) \\ &\leq \Pr(\lambda_{\min} \theta^2 \|\Delta X_1\|_{F,\min}^2 \leq 4n_r T(1 + \delta)) \\ &= \Pr(\lambda_{\min} \leq \text{SNR}^{-T} 4n_r T(1 + \delta)) \end{aligned} \quad (28)$$

$$\begin{aligned} &= \int_0^{\text{SNR}^{-T} 4n_r T(1 + \delta)} p_{\lambda_{\min}}(\lambda) d\lambda \\ &\approx p_{\lambda_{\min}}(0) \text{SNR}^{-T} 4n_r T(1 + \delta) \end{aligned} \quad (29)$$

$$\begin{aligned} &\doteq \text{SNR}^{-T} 4n_r T(1 + \delta) \\ &\doteq \text{SNR}^{-T} \end{aligned} \quad (30)$$

where in writing down (29) we have made use of the fact that the fading channel is regular. The random variable  $\|W_1\|_F^2$  that appears in the second term in (27) is a chi-squared random variable in  $2n_r T$  dimensions. Therefore

$$\begin{aligned} \Pr(\|W_1\|_F^2 \geq n_r T(1 + \delta)) &= e^{-z} \sum_{k=0}^{n_r T-1} \frac{z^k}{k!} \Big|_{z=n_r T(1 + \delta)} \\ &\doteq e^{-z} \\ &\doteq \text{SNR}^{-n_r T \beta}. \end{aligned}$$

Note that this term is independent of the channel. By choosing  $\beta > T/n_r T$ , we can make this term insignificant in comparison to  $P_1(H)$ , leading to

$$\Pr(\bar{\mathcal{A}}_1) \leq \text{SNR}^{-T}. \quad \square$$

#### APPENDIX B PROOF OF THEOREM 5

The proof makes use of the following lemma from [16]. Given a complex Hermitian matrix  $A$ , we will use  $l_k(A)$  to denote the  $k$ th ordered eigenvalue of the matrix  $A \in \mathbb{C}^{n \times n}$ , with the eigenvalues arranged in increasing order, i.e.,

$$l_1(A) \leq l_2(A) \leq \dots \leq l_n(A).$$

*Lemma 10 [16]:* Let  $A \in \mathbb{C}^{n_t \times n_t}$  be a Hermitian matrix,  $T$  be an integer such that  $1 \leq T \leq n_t$ , and  $A_T$  denote any  $T$ -by- $T$  principal submatrix of  $A$  (obtained by deleting  $n_t - T$  rows and the corresponding columns from  $A$ ). For each integer  $k$  such that  $1 \leq k \leq T$ , we have

$$l_k(A) \leq l_k(A_T) \leq l_{k+n_t-T}(A).$$

*Proof (of Theorem 5):* The optimality of  $\mathcal{X}_{\text{ARQ}}$  will be established by showing that it satisfies the Frobenius-norm, D-MG optimality and error-probability criteria of Theorem 4.

*D-MG Optimality Criterion:* It is evident from the nature of construction of the  $\mathcal{X}_{\text{ARQ}}$  code that this criterion is satisfied.

*Frobenius Norm Criterion:* Let  $\theta X \in \mathcal{X}_{\text{ARQ},L}$  and let  $\theta X_1 \in \mathcal{X}_{\text{ARQ},1}$  be the single-round ST code matrix corresponding to  $\theta X$  i.e.,  $\theta X$  is of the form

$$\theta X = \theta [X_1 \quad Z], \quad \text{for some particular } Z \in \mathbb{C}^{n_t \times T(L-1)}.$$

Let  $\theta\Delta X$  denote the difference of any two distinct matrices drawn from  $\mathcal{X}_{\text{ARQ},L}$  and  $\theta\Delta X_1$  denote the corresponding difference matrix associated with the single-round ST code. From the property of CDA-based ST codes given in (20) with  $n_t = TL$  and  $r = r_1/L$ , we obtain

$$\min_{\Delta X} \det(\theta^2 \Delta X^\dagger \Delta X) \doteq \text{SNR}^{TL-r_1/L}. \quad (31)$$

We can write

$$\Delta X^\dagger \Delta X = \begin{bmatrix} \Delta X_1^\dagger \Delta X_1 & \Delta X_1^\dagger \Delta Z \\ \Delta Z^\dagger \Delta X_1 & \Delta Z^\dagger \Delta Z \end{bmatrix}.$$

Let  $l_{\max}$  be the maximum eigenvalue of  $(\theta^2 \Delta X \Delta X^\dagger)$ . Then, we have  $l_{\max} \leq \|\theta\Delta X\|_F^2 \leq \text{SNR}$  from the energy constraint on the ST code. Using Lemma 10, we obtain

$$\begin{aligned} \min_{\Delta X_1} \det(\theta^2 \Delta X_1^\dagger \Delta X_1) &\geq \frac{\min_{\Delta X} \det(\theta^2 \Delta X \Delta X^\dagger)}{(l_{\max})^{n_t-T}} \\ &\geq \text{SNR}^{T-r_1/L}. \end{aligned} \quad (32)$$

By the arithmetic-mean geometric-mean (AM-GM) inequality, we obtain

$$\begin{aligned} \|\theta\Delta X_1\|_{\min}^2 &\geq [\det(\theta^2 \Delta X_1^\dagger \Delta X_1)]^{1/T} \\ &\geq \text{SNR}^{1-r_1/TL} \\ &= \text{SNR}^{1-r_1/n_t} \end{aligned}$$

since  $T = n_t/L$ . It follows then that  $\mathcal{T} > 0$  for all  $0 \leq r_1 < n_t$  and the Frobenius-norm criterion of Theorem 4 is thus satisfied.

*Error-Probability Criterion:* It is enough to show from Theorem 4 that

$$P_{e,l}(r_l) \leq P_{e,L}(r_L). \quad (33)$$

From the definition of the bounded distance decoder  $\mathcal{D}_l$ , it follows that  $P_{e,l}(r_l)$  is the probability of the event that

- $\theta H[\mathbf{X}_l]$  is not included in a sphere of squared radius  $n_r T l(1 + \delta)$  centered around the received vector  $\mathbf{Y}_l$ ; and
- $\theta H[\hat{\mathbf{X}}_l]$  is the unique matrix included in the sphere for some erroneous codeword  $\theta[\hat{\mathbf{X}}_l]$ .

Thus,  $P_{e,l}(r_l)$  can be upper-bounded by the probability that  $\theta[\mathbf{X}_l]$  was transmitted and the additive noise was such that  $\theta H[\mathbf{X}_l]$  is not in the sphere of the corresponding received matrix  $\mathbf{Y}_l$  (see Fig. 3). We thus have

$$P_{e,l}(r_l) \leq P_r(\|\mathbf{W}_l\|_F^2 \geq n_r T l(1 + \delta))$$

where  $\|\mathbf{W}_l\|_F^2$  is a chi-squared random variable with  $2n_r T l$  degrees of freedom. Note that this probability applies irrespective of the particular ST code employed. With  $\delta = \beta \log(\text{SNR})$  as before, we obtain

$$\begin{aligned} P_{e,l}(r_l) &\leq e^{-z} \sum_{k=0}^{n_r T l - 1} \frac{z^k}{k!} \Big|_{z=n_r T l(1+\delta)} \\ &\doteq \text{SNR}^{-n_r T l \beta}, \quad l = 1, \dots, L-1 \end{aligned} \quad (34)$$

$$\leq P_{e,L}(r_L) \quad (35)$$

by choosing a large enough value of  $\beta$ . Thus, (33) is satisfied and the proof is complete.  $\square$

## APPENDIX C

### PROOF OF THEOREM 6

Once again, the proof proceeds by verifying the D-MG optimality, Frobenius-norm and error-probability criteria.

*D-MG Optimality Criterion:* This is once again evident from the nature of construction of the ST code  $\mathcal{X}_{\text{ARQ}}$ .

*Error-Probability Criterion:* This can be proven as in the proof of Theorem 5 given in Appendix B.

*Frobenius Norm Criterion:* Note that the base alphabet under consideration for the scheme  $\mathcal{X}_{\text{ARQ}}$  is  $M^2$ -point  $\mathcal{A}_{\text{QAM}}$ . From (22) it is clear that the number of independent QAM variables in each codeword matrix is  $L^2$ , therefore

$$|\mathcal{X}_{\text{ARQ}}| = (M^2)^{L^2} = \text{SNR}^{r_1 T} \Rightarrow M^2 = \text{SNR}^{r_1/L^2}$$

as  $T = 1$  here. Let  $\theta\Delta X_1$  denote the difference matrix of any two distinct codewords in  $\mathcal{X}_{\text{ARQ},1}$ . The entries of  $\Delta X_1$  lie in  $\mathcal{O}_{\mathbb{L}'}$  which is integral over  $\mathbb{Z}[i]$ . As a result, the norm  $N_{\mathbb{L}'/\mathbb{Q}(i)}(\ell')$  of a typical entry

$$\ell' = \sum_{i=1}^k c_i \alpha_i, \quad c_i \in \mathcal{O}_{\mathbb{L}}$$

in  $\Delta X_1$  lies in  $\mathbb{Z}[i]$  and is hence lower-bounded by 1, i.e., from (14) we have

$$\prod_{j=1}^{kL} \sigma_j(\ell') \geq \text{SNR}^0. \quad (36)$$

Note that elementary column operations do not change the SNR exponent of energy of the ST code matrix. Hence from (19) and (36), we have

$$|\ell'|^2 \geq \frac{\text{SNR}^0}{(M^2)^{kL-1}}.$$

This leads to

$$\|\Delta X_1\|_F^2 \geq |\ell'|^2 \geq (M^2)^{1-kL}.$$

From (3) and the fact that every element in either the maximal subfield  $\mathbb{L}$  or else the extension field  $\mathbb{L}'$  of  $\mathbb{L}$  has magnitude with SNR exponent upper-bounded by that of  $M^2$  (see (19)) it follows that the scaling factor

$$\theta^2 \doteq \frac{\text{SNR}}{\mathbb{E}(\|\mathbf{X}_1\|_F^2)} \doteq \frac{\text{SNR}}{M^2}.$$

Therefore

$$\begin{aligned} \theta^2 \|\Delta X_1\|_F^2 &\geq \text{SNR}^{1-r_1 k/L} \\ &= \text{SNR}^{1-r_1/n_t} \\ &\Rightarrow \mathcal{T} = 1 - \frac{r_1}{n_t}. \end{aligned}$$

Therefore,  $\mathcal{T} > 0$  for all  $0 \leq r_1 < n_t$ .  $\square$

## APPENDIX D

### PROOF OF THEOREM 8

Constructions presented for the case when  $n_t|L$  have  $T = 1$ , as a result of which they are delay optimal. For the case when  $L|n_t$ , the block length of our construction is  $T = n_t/L$ . To show that  $T = n_t/L$  is the minimum possible delay, let us consider a  $n_t \times TL$  ARQ signaling ST code  $\mathcal{X}_{\text{ARQ}}$  with  $T < n_t/L$ . Let  $\theta\Delta X_L$  denote the difference matrix of any two distinct codewords from the ST code  $\mathcal{X}_{\text{ARQ},L}$ . If  $\{l_i\}_{i=1}^L$  are the nonzero

eigenvalues of  $\theta^2 \Delta X_L \Delta X_L^\dagger$ , then the pairwise error probability (PEP) [20], [21] is given as

$$\text{PEP}(\theta \Delta X_L) \doteq \left[ \frac{1}{\prod_{i=1}^{\nu} l_i} \right]^{n_r} \geq \frac{1}{(\sum_{i=1}^{\nu} l_i)^{n_r \nu}} \quad (37)$$

using the arithmetic–geometric mean inequality.

From (3), we have the energy constraint

$$\|\theta X_L\|_F^2 \leq \text{SNR}.$$

This in turn implies that

$$\|\theta \Delta X_L\|_F^2 = \sum_{i=1}^{\nu} l_i \leq \text{SNR}.$$

Substituting this value in (37), we obtain

$$\text{PEP}(\theta \Delta X_L) \geq \text{SNR}^{-n_r \nu}. \quad (38)$$

Since the above lower bound on PEP is independent of  $\Delta X_L$ , using (13) we can lower-bound the average codeword error probability  $P_e(r)$  as

$$P_e(r) \geq P_{e,L}(r_L) \geq \text{SNR}^{-n_r \nu}.$$

Note that the maximum value that  $\nu$  can attain is  $TL < n_t$ . Therefore,  $\mathcal{X}_{\text{ARQ}}$  is strictly suboptimal on the DMD tradeoff in the region  $r \approx 0$ . Hence, the constructions provided in the present paper with  $T = \lceil n_t/L \rceil$  are of minimum delay.  $\square$

#### APPENDIX E PROOF OF THEOREM 9

We prove the theorem by showing that the sufficient conditions 1), 2), and 3) spelled out in Theorem 2 hold.

Note that under the algorithm adopted for this construction, the event  $\bar{\mathcal{A}}_1$  corresponds to the event that the channel matrix  $H$  is in outage for rate  $r_1$  and hence we have that

$$\Pr(\bar{\mathcal{A}}_1) \doteq \text{SNR}^{-d_{\text{out}}(r_1)}$$

where  $d_{\text{out}}(r_1)$  is the outage exponent of the corresponding fading channel. Hence, the desired condition 1) is satisfied with  $\mathcal{T} = d_{\text{out}}(r_1)$ .

We will now show that the codes  $\mathcal{X}_{\text{ARQ},l}$ ,  $1 \leq l \leq L$  all have the property that when the channel is not in outage for rate  $r_l = r_1/l$ , that the probability of decoding error is negligible, i.e., is of order  $< \text{SNR}^{-k}$  for any integer  $k$ . Given this, it follows that

$$\begin{aligned} P_{e,l}(r_l) &\doteq \text{SNR}^{-\infty} \\ P_{e,L}(r_L) &\doteq \text{SNR}^{-d_{\text{out}}(r_L)}. \end{aligned}$$

This proves the two remaining conditions 2) and 3) as it simultaneously shows that the error probability of any intermediate decoder is less than that of the final decoder (condition 3)) and that the final decoder is a D-MG optimal code for multiplexing gain  $r_L$  (condition 2)).

#### A. Error Probability When Not in Outage

We follow the approach in [15] here. We consider the PEP of the decoder  $\mathcal{D}_l$ . We have

$$\begin{aligned} d_E^2 &= \sum_{k=1}^l \theta^2 \|H \Delta \phi^{k-1}(X)\|^2 \\ &= \theta^2 \text{Tr}(\mathbf{H}_l^\dagger \mathbf{H}_l \Delta X_l \Delta X_l^\dagger) \end{aligned}$$

where, for  $1 \leq l \leq m$ , we will use  $\mathbf{H}_l$ ,  $\Delta X_l$  to denote the block-diagonal matrices

$$\begin{aligned} \mathbf{H}_l &= \begin{bmatrix} H & & & \\ & H & & \\ & & \ddots & \\ & & & H \end{bmatrix}_{n_r l \times n l} \\ \Delta X_l &= \begin{bmatrix} \Delta X & & & \\ & \phi(\Delta X) & & \\ & & \ddots & \\ & & & \phi^{l-1}(\Delta X) \end{bmatrix}_{n l \times n l} \end{aligned}$$

Here again, while the codes used in the ARQ scheme correspond to  $l$  in the range  $1 \leq l \leq L$ , we extend the definition to include all  $l$  in the range  $1 \leq l \leq m$  as this will be found useful in the proofs to follow. Let

$$\lambda_1 \leq \lambda_2 \leq \dots \leq \lambda_{ln} \quad (39)$$

$$\mu_1 \geq \mu_2 \geq \dots \geq \mu_{ln} \quad (40)$$

be an ordering of the eigenvalues of  $\mathbf{H}_l^\dagger \mathbf{H}_l$ ,  $\Delta X_l \Delta X_l^\dagger$ , respectively, and let  $\{\alpha_i\}$  be defined as before by

$$\lambda_i = \text{SNR}^{-\alpha_i}, \quad 1 \leq i \leq ln.$$

Using the mismatched eigenvalue bound [10], [22], we obtain

$$d_E^2 \geq \theta^2 \sum_{i=1}^{ln} \lambda_i \mu_i.$$

Let  $\{\nu_j\}_{j=1}^{mn}$  denote the eigenvalues of  $\Delta X_m \Delta X_m^\dagger$ . Since the eigenvalues of  $\Delta X_l \Delta X_l^\dagger$  are a subset of the eigenvalues of  $\Delta X_m \Delta X_m^\dagger$ , we assume, without loss of generality, an ordering of the eigenvalues  $\{\nu_j\}$  such that

$$\nu_j = \mu_j, \quad 1 \leq j \leq ln.$$

Then for every  $1 \leq J \leq ln$ , we have

$$\begin{aligned} d_E^2 &\geq \theta^2 \sum_{i=ln-J}^{ln} \lambda_i \mu_i \\ &= \theta^2 \sum_{i=ln-J}^{ln} \lambda_i \nu_i \\ &\geq \theta^2 \left( \prod_{i=ln-J}^{ln} \lambda_i \right)^{1/J+1} \left( \prod_{i=ln-J}^{ln} \nu_i \right)^{1/J+1} \end{aligned}$$

$$\begin{aligned}
&= \theta^2 \left( \prod_{i=ln-J}^{ln} \lambda_i \right)^{1/J+1} \\
&\quad \times \left\{ \frac{\prod_{i=1}^{mn} (\nu_i)}{\prod_{i=1}^{ln-J-1} (\nu_i) \prod_{i=ln+1}^{mn} (\nu_i)} \right\}^{1/J+1} \\
&\geq \theta^2 \left( \frac{\text{SNR}^{-\sum_{i=ln-J}^{ln} \alpha_i}}{(M^2)^{mn-(J+1)}} \right)^{1/J+1} \doteq \text{SNR}^{\delta_J/J+1}
\end{aligned}$$

where

$$\begin{aligned}
\delta_J &= (J+1) \left( 1 - \frac{r_1}{mn} \right) - (mn - (J+1)) \frac{r_1}{mn} \\
&\quad - \sum_{i=ln-J}^{ln} \alpha_i = J+1 - r_1 - \sum_{i=ln-J}^{ln} \alpha_i \\
&= \sum_{i=ln-J}^{ln} (1 - \alpha_i) - r_1.
\end{aligned}$$

In this derivation we have made use of the fact that the product of the eigenvalues is equal to the determinant, the nonvanishing determinant property enunciated in (23), the fact that the eigenvalues of a matrix are upper-bounded by the trace, and (25), (26). We will now show that if the block-fading channel is not in outage for rate  $r_1/l$ , that for some  $J$ ,  $1 \leq J \leq ln$ ,  $\delta_J > 0$ . If the block-fading channel is not in outage for rate  $r_1 + \epsilon/l$ , we must have

$$\begin{aligned}
\log \det(I_n + \text{SNR}H^\dagger H) &\geq \frac{r_1 + \epsilon}{l} \log(\text{SNR}) \\
\Rightarrow \log \det(I_{ln} + \text{SNR}\mathbf{H}_l^\dagger \mathbf{H}_l) &\geq (r_1 + \epsilon) \log(\text{SNR}) \\
&\Rightarrow \sum_{i=1}^{ln} (1 - \alpha_i)^+ \geq (r_1 + \epsilon).
\end{aligned}$$

Let  $ln - J$  be the smallest index  $k$  of  $\alpha$  for which  $\alpha_k < 1$ . Then this is equivalent to the condition

$$\begin{aligned}
\sum_{i=ln-J}^{ln} (1 - \alpha_i) &\geq (r_1 + \epsilon) \\
\Rightarrow (J+1) - \sum_{i=ln-J}^{ln} \alpha_i &\geq (r_1 + \epsilon) \\
\Rightarrow \delta_J &\geq \epsilon > 0.
\end{aligned}$$

By taking the limit as  $\epsilon \rightarrow 0$  we see as desired, that the probability of error is negligible in the no-outage region. Clearly, this property bestows upon the ARQ code constructed here, the property of being approximately universal.

## REFERENCES

- [1] H. El Gamal, G. Caire, and M. O. Damen, "The MIMO ARQ channel: Diversity-multiplexing-delay tradeoff," *IEEE Trans. Inf. Theory*, vol. 52, no. 8, pp. 3601–3621, Aug. 2006.
- [2] R. Heath and A. Paulraj, "Switching between multiplexing and diversity based on constellation distance," in *Proc. Allerton Conf. Communications, Control and Computing*, Monticello, IL, Oct. 2000.

- [3] L. Zheng and D. N. C. Tse, "Diversity and multiplexing: A fundamental tradeoff in multiple-antenna channels," *IEEE Trans. Info. Theory*, vol. 49, no. 5, pp. 1073–1096, May 2003.
- [4] P. Dayal and M. K. Varanasi, "An optimal two transmit antenna space-time code and its stacked extensions," in *Proc. Asilomar Conf. Signals, Systems and Computers*, Monterey, CA, Nov. 2003.
- [5] H. Yao and G. W. Wornell, "Structured space-time block codes with optimal diversity-multiplexing tradeoff and minimum delay," in *Proc. IEEE GLOBECOM 2003*, San Francisco, CA, Dec. 2003, vol. 4, pp. 1941–1945.
- [6] J.-C. Belfiore, G. Rekaya, and E. Viterbo, "The golden code: A  $2 \times 2$  full-rate space-time code with non-vanishing determinants," in *Proc. IEEE Int. Symp. Information Theory (ISIT 2004)*, Chicago, IL, Jun. 2, 2004, p. 308.
- [7] H. El Gamal, G. Caire, and M. O. Damen, "Lattice coding and decoding achieve the optimal diversity-multiplexing tradeoff of MIMO channels," *IEEE Trans. Inf. Theory*, vol. 50, no. 6, pp. 968–985, Jun. 2004.
- [8] S. Tavildar and P. Viswanath, "Approximately universal codes over slow-fading channels," *IEEE Trans. Inf. Theory*, vol. 52, no. 7, pp. 3233–3258, Jul. 2006.
- [9] Kiran. T. and B. S. Rajan, "STBC-schemes with non-vanishing determinant for certain number of transmit antennas," *IEEE Trans. Inf. Theory*, vol. 51, no. 8, pp. 2984–2992, Aug. 2005.
- [10] P. Elia, K. R. Kumar, S. A. Pawar, P. V. Kumar, and H.-F. Lu, "Explicit, minimum-delay space-time codes achieving the diversity-multiplexing gain tradeoff," *IEEE Trans. Inf. Theory*, vol. 52, no. 9, pp. 3869–3884, Sep. 2006.
- [11] F. Oggier, G. Rekaya, J. C. Belfiore, and E. Viterbo, "Perfect space-time block codes," *IEEE Trans. Inf. Theory*, vol. 52, no. 9, pp. 3885–3902, Sep. 2006.
- [12] P. Elia, B. A. Sethuraman, and P. V. Kumar, "Perfect space-time codes with minimum and non-minimum delay for any number of antennas," *IEEE Trans. Inf. Theory*, vol. 53, no. 11, pp. 3853–3868, Nov. 2007.
- [13] B. A. Sethuraman and B. S. Rajan, "Full-rank, full-rate STBCs from division algebras," in *Proc. 2002 IEEE Information Theory Workshop (ITW 2002)*, Bangalore, India, Oct. 2002, pp. 69–72.
- [14] B. A. Sethuraman, B. S. Rajan, and V. Shashidhar, "Full-diversity, high-rate, space-time block codes from division algebras," *IEEE Trans. Inf. Theory*, vol. 49, no. 10, pp. 2596–2616, Oct. 2003.
- [15] P. Elia and P. V. Kumar, "Approximately-Universal Space-Time Codes for the Parallel Multi-Block and Cooperative-Dynamic Decode-and-Forward Channels," 2007, preprint, arXiv:0706.3502v2 cs.IT.
- [16] R. A. Horn and C. R. Johnson, *Matrix Analysis*. Cambridge, U.K.: Cambridge Univ. Press, 1999.
- [17] A. Edelman, "Eigen values and condition numbers of random matrices," *SIAM J. Matrix Anal. Appl.*, vol. 9, no. 4, pp. 543–560, Oct. 1988.
- [18] A. A. Albert, *Structure of Algebras*. Providence, RI: Amer. Math. Soc., 1961, Coll. Publ., vol. 24.
- [19] J.-C. Belfiore and G. Rekaya, "Quaternionic lattices for space-time coding," in *Proc. IEEE Information Theory Workshop.*, Paris, France, Apr. 2003, pp. 267–270.
- [20] 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," in *Proc. IEEE Vehicular Technology Conf. (VTC'96)*, Atlanta, GA, Apr./May 1996, pp. 136–140.
- [21] V. Tarokh, N. Seshadri, and A. R. Calderbank, "Space-time codes for high data rate wireless communication: Performance criterion and code construction," *IEEE Trans. Inf. Theory*, vol. 44, no. 2, pp. 744–765, Mar. 1998.
- [22] C. Köse and R. D. Wesel, "Universal space-time trellis codes," *IEEE Trans. Inf. Theory*, vol. 49, no. 10, pp. 2717–2727, Oct. 2003.
- [23] S. Yang, J.-C. Belfiore, and G. R.-B. Othman, "Perfect space-time block codes for parallel MIMO channels," in *Proc. IEEE Int. Symp. Information Theory (ISIT)*, Seattle, WA, Jul. 2006, pp. 1949–1953.
- [24] H.-F. Lu, "Explicit construction of multi-block space-time codes that achieve the diversity-multiplexing gain tradeoff," in *Proc. IEEE Int. Symp. Information Theory (ISIT)*, Seattle, WA, Jul. 2006, pp. 1149–1153.

**Sameer A. Pawar** received the B.E. degree from Government College of Engineering, Pune, India, in 2001 and the M.Sc.(Engg.) degree from the Indian Institute of Science, Bangalore, in 2005.

He is currently working toward the Ph.D. degree in the Department of Electrical Engineering and Computer Science, University of California, Berkeley.

His current research interests include MIMO systems, cooperative communications, Network Coding and Information theory.

**K. Raj Kumar** (S'02) received the B.E. degree from the University of Madras, Madras, India, in 2003 and the M.Sc.(Engg.) degree from the Indian Institute of Science, Bangalore, in 2005.

He is currently working toward the Ph.D. degree in the Department of Electrical Engineering–Systems, University of Southern California (USC), Los Angeles. His current research interests include MIMO systems, cooperative communications, cognitive radios, and multiuser information theory.

Mr. Kumar is a recipient of the 2006 Best Student Paper Award from the Department of Electrical Engineering–Systems, USC, and an Oakley Fellowship from the Graduate School at USC for the 2007–2008 academic year.

**Petros Elia** received the B.Sc. degree in electrical engineering from the Illinois Institute of Technology, Chicago, in 1997. In 2001 and 2006, respectively, he received the M.Sc. and Ph.D. degrees in electrical engineering from the University of Southern California, Los Angeles.

He is currently Assistant Professor within the Department of Mobile Communications at EURECOM, Sophia-Antipolis, France. His research interests include information-theoretic and coding aspects of wireless communications, cooperative communications, complexity theory, and cross-layer optimization.

**P. Vijay Kumar** (S'80-M'82-SM'01-F'02) received the B.Tech. and M.Tech. degrees from Indian Institute of Technology, Kharagpur, and Indian Institute of

Technology, Kanpur, respectively, and the Ph.D. degree from the University of Southern California (USC), Los Angeles, in 1983, all in electrical engineering.

From 1983 to 2003, he was on the faculty of the Department of Electrical Engineering–Systems at USC and since 2003, he has been a Professor at the Electrical and Computer Engineering Department of the Indian Institute of Science, Bangalore, India, on leave of absence from USC.

From 1993–1996, Prof. Kumar was an Associate Editor for Coding Theory for the IEEE TRANSACTIONS ON INFORMATION THEORY. He is corecipient of the 1995 IEEE Information Theory Society's Prize Paper Award as well as the 1994 USC School-of-Engineering Senior Research Award for contributions to coding theory. A family of low-correlation sequences introduced in a 1996 paper coauthored by him is part of the 3G, WCDMA, mobile communication standard. He is also corecipient of a best paper award given at the 2008 IEEE International Conference on Distributed Computing in Sensor Systems.

**B. A. Sethuraman** received the bachelor's degree in mechanical engineering from the Indian Institute of Technology, Chennai, India. He then switched fields and received the Ph.D. degree in mathematics from the University of California, San Diego, La Jolla.

He is with the Department of Mathematics at California State University, Northridge. His research interests are in algebra and algebraic geometry, in which he has published extensively, won several research grants, and written textbooks. Recently, he has helped to develop new applications of algebra to wireless communication in collaboration with engineers; in particular, his suggestion that division algebras are the correct mathematical objects for use in multiple-input multiple-output (MIMO) applications has firmly taken root, and codes based on division algebras now come close to achieving the fundamental limits of outage-limited communications.