Diversity Embedded Codes: Theory and Practice

S. Dusad S. N. Diggavi N. Al-Dhahir A. R. Calderbank

Abstract

Diversity embedded codes are high-rate space-time codes that have a high-diversity code embedded within them. They allow a form of communication where the high-rate code opportunistically takes advantage of good channel realizations while the embedded high-diversity code provides guarantees that at least part of the information is received reliably. Over the past few years, code designs and fundamental limits of performance for such codes have been developed. In this paper, we review these ideas by giving the developments in a unified framework. In particular, we present both the coding technique as well as information-theoretic bounds in the context of Intersymbol Interference (ISI) channels. We investigate the systems implications of diversity embedded codes by examining value to network utility maximization, unequal error protection for wireless transmission, rate opportunism and packet delay optimization.

I. Introduction

Over the past decade, since the seminal work of [40], [25], [42], multiple -antenna (space-time) codes have emerged as a means of enabling reliable high-data rate wireless communications. There have been significant developments both in the theory and practice of space-time codes over the past few years. These have led to adoption of space-time codes in next-generation wireless communication standards such as 802.11n and WiMAX. There is an inherent tension between rate and reliability exemplified in space-time codes. This was explored in the context of finite-rate (fixed alphabet size) codes in [42] and in the context of information-theoretic rate growth in [45]. Both of these results show that to achieve a high transmission rate, one might need to sacrifice reliability (diversity) and vice-versa. Therefore, the main question addressed by most researchers has been on how to design transmission techniques that achieve a particular point on the rate-reliability tradeoff curve. For example, the emphasis in the coding literature has been the design of maximal-diversity codes, *i.e.* one extremal point in this tradeoff. In [40],

EPFL, Lausanne, Switzerland, Email: {sanket.dusad,suhas.diggavi}@epfl.ch. S. Dusad was supported in part by SNSF Grant # 200021-105640/1. S. N. Diggavi is part of the SNSF supported NCCR-MICS center on wireless sensor networks.

UT, Dallas, Email: aldhahir@utdallas.edu. The work of N. Al-Dhahir is supported in part by NSF contracts CCF 0430654 and DMS 0528010.

Princeton University, Email: calderbank@math.princeton.edu. A R. Calderbank was supported in part by NSF grant # 1096066.

[25], the focus is on obtaining the maximal spatial multiplexing rate for ergodic channels. There are also codes which achieve particular points on the rate-reliability tradeoff curve (see for example, [35] and references therein for a history of such codes).

Diversity embedded codes introduced in [12], [13] ask a *different* question where multiple levels of reliabilities (in terms of error probability or diversity order) were sought for different messages. Since then, there have been (algebraic) multi-level constructions for diversity embedded codes [17], [23], an understanding of fundamental limits of such codes [15], [16], [22] as well as its impact on network layer protocols [32], [19]. This paper summarizes these ideas in a unified framework.

One motivation for providing such a physical layer characteristic is the need for wireless networks to support a variety of applications with different quality-of-service (QoS) requirements. For example, real-time (multimedia) applications need lower delay and therefore higher reliability (diversity) as compared to non-real-time applications. In particular, if we design the overall system for a fixed rate-diversity operating point, we might be over-provisioning a resource which could be flexibly allocated to different applications. This thought process leads to viewing diversity (reliability) as a systems resource that can be allocated judiciously to satisfy the QoS requirements for the different applications. This is the characteristic of diversity embedded codes which can also be thought of as unequal error protection (UEP) codes designed for wireless fading channels. The idea of UEP has a long and rich history (see for example [39], [8] and references therein). UEP codes have been designed for the binary Hamming distance metric (see for example [36]) and for the Euclidean distance metric encountered in the Gaussian channel (see for example [5]). Hence, diversity embedded codes provide UEP with respect to the diversity metric suitable for fading channels [17].

In transmitting over an unknown channel, whose capacity is therefore not known a priori, one can envisage two strategies. One strategy would be conservative where we design for the worst channel for a given reliability (outage). Another strategy would be opportunistic where we embed information in such a manner so as to deliver part of the information if the channel is bad, but more information when the channel is good. It is clear that one would do better with a conservative strategy in an adversarial situation, since it would ensure a higher rate in the worst-case. However, wireless channels being random (as opposed to adversarial), we can opportunistically gain rate in the latter strategy. Therefore, another interpretation of diversity embedded codes is in terms of opportunistic communications, where one takes advantage of benign channels while giving some guarantees when deep fades occur.

In this survey of diversity embedded codes, we focus our attention on designs which are appropriate for broadband inter-symbol interference (ISI) channels. However, the designs specialize to the flat-fading case, where sometimes a simpler code would actually suffice. This simplification will be pointed out in Sections IV and V. One important point to note is that if one wants to take advantage of the multipath diversity offered by ISI channels, we *need to design* codes specifically for ISI channels. If one uses space-time codes suitable for flat fading channels, we cannot guarantee in general that these codes would be capable of gathering more than the spatial diversity gains. This becomes especially important in a rich multipath environment, where we can design codes at a higher rate, and utilize the multipath (ISI) diversity to improve reliability. In fact this argument motivates the importance of diversity embedded codes for ISI channels, since such codes give flexibility in the design of the rate-diversity performances needed.

One of the main classifications for the diversity embedded codes is in terms of the constraints imposed on the design. In many applications, the need to control transmit parameters (like peak-to-average ratio) motivates a transmit alphabet constraint, where the transmit symbol is restricted to come from a fixed set (like M-QAM, M-PSK etc.). For such cases, there is a well-defined tradeoff between rate and diversity (see Section II-B). Therefore, designs of diversity embedded codes with transmit alphabet constraints is the focus of Section IV. The multi-level constructions given are natural generalizations of space-time codes [42].

If we move away from transmit alphabet constraints, the appropriate tradeoff to consider is the *diversity-multiplexing* tradeoff introduced in [45]. In Section V, we design diversity embedded codes appropriate to this regime and examine fundamental information-theoretic performance limits. In particular, it is shown that when we have a single-degree of freedom (*i.e.*, when we have either one transmit antenna or one receive antenna; SIMO/MISO), the diversity multiplexing tradeoff is *successively refinable* [15], [22]. This means that one can design diversity embedded codes such that the higher reliability stream gets optimal performance (*i.e.*, it is extremal on the diversity-multiplexing tradeoff), while the overall code (sum-rate of messages) is also optimal. This implies that (asymptotically in SNR) one can perfectly embed codes, therefore creating *ideal* opportunistic codes that automatically adjust multiplexing rate to that level supported by the channel (without a priori channel knowledge). The intuition for this result through scalar broadcast channels is illustrated in Section V.

The outline of the paper is as follows. We begin with the notation and a presentation of results on the tradeoff between rate and reliability in Section II. In Section III, we introduce the concept of diversity embedding and give design criteria for finite alphabet and rate growth codes. We present a class of linear and multilevel diversity embedded code construction for the transmit alphabet constraint case in IV. In Section V, we focus on the information-theoretic limits of diversity embedded codes and demonstrate the successive refinement property. The networking implications (including wireless multimedia delivery) and the benefits of the diversity embedded codes are illustrated in Section VI.

II. PRELIMINARIES

Our focus is on the quasi-static fading channel where we transmit information coded over M_t transmit antennas with M_r antennas at the receiver. Throughout this paper, we assume that the transmitter has no channel state information (CSI), whereas the receiver is able to perfectly track the channel (a common assumption, see for example [2], [42]). Differential decoding schemes for diversity-embedding codes which do not require receiver CSI were designed and analyzed in [37].

The coding scheme is limited to one quasi-static transmission block of large enough block size $T \geq T_{thr}$ to be specified later. The received vector at time n after demodulation and sampling can be written as

$$\mathbf{y}[n] = \mathbf{H}_0\mathbf{x}[n] + \mathbf{H}_1\mathbf{x}[n-1] + \ldots + \mathbf{H}_\nu\mathbf{x}[n-\nu] + \mathbf{z}[n] \tag{1}$$

where $\mathbf{y} \in \mathbb{C}^{M_r \times 1}$ is the received vector at time n, $\mathbf{H}_l \in \mathbb{C}^{M_r \times M_t}$ represents the l^{th} matrix tap of the MIMO ISI channel, $\mathbf{x}[n] \in \mathbb{C}^{M_t \times 1}$ is the space-time coded transmission vector at time n with transmit power constraint P and $\mathbf{z} \in \mathbb{C}^{M_r \times 1}$ is assumed to be additive white (temporally and spatially) Gaussian noise with variance σ^2. The matrix $\mathbf{H}_l$ consists of fading coefficients h_{ij} which are i.i.d. $\mathcal{CN}(0,1)$ and fixed for the duration of the block length (T). For the special case of a flat-fading channel, assuming $\nu = 0$ the received symbols over a block of length T can be written as

$$\mathbf{Y} = \mathbf{H}\mathbf{X} + \mathbf{Z}, \tag{2}$$

where $\mathbf{Y} = [\mathbf{y}(0), \ldots, \mathbf{y}(T-1)] \in \mathbb{C}^{M_r \times T}$ and $\mathbf{X} = [\mathbf{x}(0), \ldots, \mathbf{x}(T-1)] \in \mathbb{C}^{M_t \times T}$ is the space-time-coded transmission sequence.

For the channel models in (1), (2) define the notion of diversity order [42] as follows,

Definition 2.1: A coding scheme which has an average error probability $P_e(\text{SNR})$ as a function of SNR that behaves as

$$\lim_{\text{SNR}\to\infty} \frac{\log(P_e(\text{SNR}))}{\log(\text{SNR})} = -D, \tag{3}$$

is said to have a diversity order of D.

In words, a scheme with diversity order D has an error probability at high SNR behaving as $P_e(\text{SNR}) \approx \text{SNR}^{-D}$. We use the special symbol $\doteq$ to denote exponential equality *i.e.*, we write $f(SNR) \doteq SNR^b$ to denote

$$\lim_{SNR\to\infty} \frac{\log f(SNR)}{\log(SNR)} = b$$

and $\dot{\leq}$ and $\dot{\geq}$ are defined similarly.

A. Review of Space-Time Code Design Criteria

For codes designed for a finite (and fixed) rate, one can bound the error probability using *pairwise error probability* (PEP) between two candidate codewords. This leads to the now well-known rank criterion for determining the diversity order of a space-time code [30], [42]. Considering a codeword sequence $\mathbf{X}$ as defined in (2), the PEP between two codewords $\mathbf{X}$ and $\mathbf{E}$ can be determined by the codeword difference matrix $\mathbf{B}(\mathbf{X},\mathbf{E}) = \mathbf{X} - \mathbf{E}$ [30], [42] as follows

$$P_e(\text{SNR}, \mathbf{X} \longrightarrow \mathbf{E}) \doteq \text{SNR}^{-M_r \text{rank}(\mathbf{B}(\mathbf{X},\mathbf{E}))}. \tag{4}$$

For fixed-rate space time codes, by using the simple union bound argument, it can be shown that the diversity order is given by [42] as,

$$D = M_r \min_{\mathbf{X} \neq \mathbf{E} \in \mathcal{C}} \text{rank}(\mathbf{B}(\mathbf{X},\mathbf{E})) \ . \tag{5}$$

The error probability is determined by *both* the coding gain and the diversity order. Hence, the code design criterion prescribed in [42] is to design the codebook $\mathcal{C}$ so that the minimal rank of the codeword difference matrix corresponds to the required diversity order and the minimal determinant gives the corresponding coding gain. For codes designed for multiplexing rate an alternative design criterion [24] which bounds the minimum determinant of the difference matrix is more appropriate. Assuming $T \geq M_t$ and a communication rate of SNR^r the criterion is given by,

$$\min_{\mathbf{X} \neq \mathbf{E} \in \mathcal{C}} \det(\mathbf{B}(\mathbf{X},\mathbf{E})) \dot{\geq} SNR^{M_t - r} \tag{6}$$

and is referred to as the Non Vanishing Determinant (NVD) criterion.

B. Tradeoff between Rate and Reliability

For a given diversity order, it is natural to ask for upper bounds on the achievable rate. For a flat Rayleigh fading channel with a fixed transmit alphabet constraint, the tradeoff has been examined in [42] where the following result was obtained.

Theorem 2.2: **(Transmit-alphabet constrained, rate-diversity tradeoff for flat-fading channels)** [34], [42] If a constellation of size $|\mathcal{A}|$ is used for transmission and the diversity order of the system is qM_r, then the rate R that can be achieved is bounded as

$$R \leq (M_t - q + 1) \log_2 |\mathcal{A}| \tag{7}$$

in bits per transmission.

Another point-of-view explored in [45], allows the rate of the codebook to increase with SNR, *i.e.*, defines a multiplexing rate. If we consider a sequence of coding schemes with transmission rate as a function of SNR given by $R(SNR)$, then the multiplexing rate r is defined as

$$r = \lim_{SNR \to \infty} \frac{R(SNR)}{\log(SNR)}. \tag{8}$$

Therefore, just as Theorem 2.2 shows the tension between achieving high-rate and high-diversity for a fixed-transmit-alphabet constraint, there exists a tension between multiplexing rate and diversity as well [45], a special case of which is described next.

Theorem 2.3: **(Diversity-Multiplexing (D-M) tradeoff for flat-fading channels)** [45] The diversity-multiplexing tradeoff for flat-fading channels with a single degree of freedom *i.e.*, $\min(M_t, M_r) = 1$, is given by

$$D(r) = max(M_t, M_r)(1 - r) \tag{9}$$

for $0 \le r \le 1$.

Similarly, the D-M tradeoff for transmission over a SISO ISI channel was established in [29].

Theorem 2.4: **(D-M tradeoff for SISO ISI channels)** [29] The diversity multiplexing tradeoff for transmission over a SISO ISI channel with $\nu + 1$ taps for transmission over a period of time T is bounded by

$$(\nu + 1)\left(1 - \frac{T}{T - \nu} r\right) \le D_{isi}(r) \le (\nu + 1)(1 - r) \tag{10}$$

for $0 \le r \le \frac{T-\nu}{T}$.

However, the D-M trade-off for multiple antenna ISI channels is still not known in general. One of the consequences of the results in Section V is to establish the trade-off for the MISO/SIMO cases.

III. Diversity Embedding

The classical approach towards code design for channels is to maximize the data rate given a desired level of reliability. A natural setting to address this question is the outage formulation. The classical outage formulation divides the set of channel realizations into an outage set $\mathcal{O}$ and a non-outage set $\overline{\mathcal{O}}$: it requires that a code has to be designed such that the transmitted message can be decoded with arbitrary small error probability on all the channels in the non-outage set. Since the code must work for all such channels, the data rate is limited by the worst channel in the non-outage set. Note that in this scenario, the communication strategy cannot take advantage of the opportunity when the channel happens to be stronger than the worst channel in the non-outage set.

Diversity embedded coding takes advantage of the good channel realizations by an opportunistic coding strategy. Consider two streams of messages $\mathcal{H}$, the high priority message stream, and $\mathcal{L}$, the low priority message stream. Diversity embedded codes encode the streams such that the high-priority stream ($\mathcal{H}$) is decoded with arbitrary small error probability whenever the channel is not in outage ($\mathcal{O}$) and in addition the lower-priority stream is decoded whenever the channel is in a set $\mathcal{G} \subset \overline{\mathcal{O}}_H$ of good channels (see Figure 1).

The decoder jointly decodes the two message sets and define two error probabilities, $P_e^H(SNR)$ and $P_e^L(SNR)$, which denote the average error probabilities for message sets $\mathcal{H}$ and $\mathcal{L}$, respectively. Then, analogous to Definition 2.1, the diversity order for the messages can be written as

$$D_H = \lim_{\text{SNR}\to\infty} \frac{-\log P_e^H(SNR)}{\log(\text{SNR})}, \;\; D_L = \lim_{\text{SNR}\to\infty} \frac{-\log P_e^L(SNR)}{\log(\text{SNR})}. \tag{11}$$

Denote R_H and R_L to be the supportable rates for the worst channels in $\overline{\mathcal{O}}_H$ and $\mathcal{G}$, respectively. Then, we can characterize the achievable tuple (R_H, D_H, R_L, D_L) for the cases of finite-alphabet and rate-growth codes for transmission over ISI and flat-fading channels.

A. Diversity Embedding Coding Scheme for ISI Channels

For ISI channels, we will consider a transmission scheme in which we transmit over a period $T-\nu$ and send (fixed) known symbols[1] for the last ν transmissions. For the period of communication we can equivalently write the received data as follows

$$\underbrace{\begin{bmatrix} \mathbf{y}[0] & \dots & \mathbf{y}[T-1] \end{bmatrix}}_{\mathbf{Y}} = \underbrace{\begin{bmatrix} \mathbf{H}_0 & \dots & \mathbf{H}_\nu \end{bmatrix}}_{\mathbf{H}} \underbrace{\begin{bmatrix} \mathbf{x}[0] & \mathbf{x}[1] & \dots & \mathbf{x}[T-\nu-1] & 0 & \dots & 0 \\ 0 & \mathbf{x}[0] & \mathbf{x}[1] & \dots & \mathbf{x}[T-\nu-1] & 0 & 0 \\ \dots & & \vdots & \ddots & . & . & \vdots \\ 0 & \dots & 0 & \mathbf{x}[0] & \mathbf{x}[1] & \dots & \mathbf{x}[T-\nu-1] \end{bmatrix}}_{\mathbf{X}} + \mathbf{Z} \tag{12}$$

i.e.,

$$\mathbf{Y} = \mathbf{H}\mathbf{X} + \mathbf{Z} \tag{13}$$

where $\mathbf{Y} \in \mathbb{C}^{M_r \times T}$, $\mathbf{H} \in \mathbb{C}^{M_r \times (\nu+1)M_t}$, $\mathbf{X} \in \mathbb{C}^{(\nu+1)M_t \times T}$, $\mathbf{Z} \in \mathbb{C}^{M_r \times T}$. Notice that the structure in (13) is different from the flat-fading case as the channel imposes a banded Toeplitz structure on the equivalent space-time codewords $\mathbf{X}$ given in (12). Although this structure makes the design of space-time codes different from the flat-fading case, the analysis in [30], [42] can be easily extended to fading ISI channels and analogous to (5) the diversity order for a *fixed rate transmission* over an ISI channel can be written as[2]

$$D = M_r \min_{\mathbf{X}_1 \neq \mathbf{X}_2} \text{rank}(\mathbf{X}_1 - \mathbf{X}_2). \tag{14}$$

Note that here $\mathbf{X}_1, \mathbf{X}_2 \in \mathbb{C}^{(\nu+1)M_t \times T}$ are matrices with structure given in (12). For reference, the space-time codeword $\mathbf{X}$ is completely determined by the matrix $\mathbf{X}^{(1)}$ given by

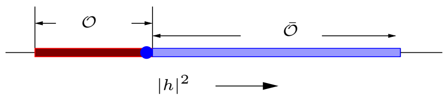
$$\mathbf{X}^{(1)} = \begin{bmatrix} \mathbf{x}[0] & \mathbf{x}[1] & \dots & \mathbf{x}[T-\nu-1] & 0 & \dots & 0 \end{bmatrix} \tag{15}$$

and define a mapping Θ_ν of a matrix of dimension $M_t \times T$ to a matrix of dimension $(\nu+1)M_t \times T$ by

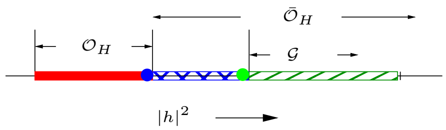
$$\Theta_\nu\left(\mathbf{X}^{(1)}\right) = \begin{bmatrix} \mathbf{x}[0] & \mathbf{x}[1] & \dots & \mathbf{x}[T-\nu-1] & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{0} & \mathbf{x}[0] & \mathbf{x}[1] & \dots & \mathbf{x}[T-\nu-1] & \mathbf{0} & \mathbf{0} \\ \dots & & \vdots & \ddots & . & . & \vdots \\ \mathbf{0} & \dots & \mathbf{0} & \mathbf{x}[\mathbf{0}] & \mathbf{x}[1] & \dots & \mathbf{x}[T-\nu-1] \end{bmatrix}. \tag{16}$$

[1]Taken without loss of generality to be zero symbols. This plays the role of isolating transmission blocks by eliminating inter block interference (due to ISI).

[2]Note that when we consider multiplexing rate, the diversity order expression would be different; see Section V.



(a) Classical setting



(b) Diversity embedded setting

Fig. 1. Outage events in the classical setting and for diversity embedded coding

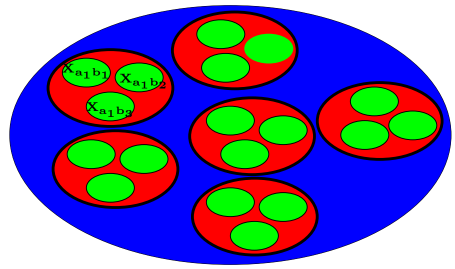


Fig. 2. Codeword Clusters

For the case of a scalar ISI channel with $(\nu+1)$ taps and a single transmit antenna, it can be shown by a simple argument (see for example [43]) that an uncoded transmission scheme can achieve a diversity order of $(\nu+1)$. However, in the multiple-transmit-antenna case, it is not obvious that a space-time code designed for a flat-fading channel can achieve such a $(\nu+1)-$fold increase in the diversity order. Therefore, the design of codes for fading ISI channels cannot be immediately done by using the codes for flat-fading channels. A finite-alphabet construction to exploit the available multipath diversity gains from ISI channels with M_t multiple transmit antennas was proposed in [28] for the maximum-diversity case but the rate of the code for this construction was $1/M_t$ as opposed to the maximal possible rate of 1, which is shown to be achievable in Section IV.

B. Design Criteria

To ensure embedding of diversity, the transmitted space time codewords $\{\mathbf{X}_k\}$ can be divided into codeword clusters as shown in Figure III-B, where the codeword clusters shown correspond to all $\{\mathbf{X}_k\}$ arising due to a particular message $\mathbf{a} \in \mathcal{H}$. The joint decoding of the two message sets yields candidate clusters and elements within the clusters. Given a desired tuple (R_H, D_H, R_L, D_L), the rates (R_H, R_L) and the design criteria depend on the transmit alphabet size as follows:

- **Finite Transmit Alphabet**: For transmission from an alphabet $\mathcal{A}$, we need to design message sets with sizes $|\mathcal{A}|^{R_H T}$ and $|\mathcal{A}|^{R_L T}$. Denote $\mathbf{X}_{\mathbf{a},\mathbf{b}}$ to be the codeword corresponding to message $\mathbf{a}$ on the stream $\mathcal{H}$ and message $\mathbf{b}$ on the stream $\mathcal{L}$. For transmission over flat-fading and ISI channels, the following code design criterion ensure that the the diversities D_H and D_L are achieved

$$\min_{\mathbf{a}_1 \neq \mathbf{a}_2 \in \mathcal{H}} \min_{\mathbf{b}_1, \mathbf{b}_2 \in \mathcal{L}} \text{rank}(\mathbf{X}_{\mathbf{a}_1,\mathbf{b}_1} - \mathbf{X}_{\mathbf{a}_2,\mathbf{b}_2}) \quad \geq D_H/M_r, \tag{17}$$

$$\min_{\mathbf{b}_1 \neq \mathbf{b}_2 \in \mathcal{L}} \min_{\mathbf{a}_1, \mathbf{a}_2 \in \mathcal{H}} \text{rank}(\mathbf{X}_{\mathbf{a}_1,\mathbf{b}_1} - \mathbf{X}_{\mathbf{a}_2,\mathbf{b}_2}) \quad \geq D_L/M_r. \tag{18}$$

- **Rate Growth**: Let $P_H(SNR)$ and $P_L(SNR)$, denote the average error probabilities for message sets $\mathcal{H}$ and $\mathcal{L}$ respectively. For each 4-tuple (R_H, D_H, R_L, D_L), define a corresponding (r_H, D_H, r_L, D_L)

as an achievable rate-diversity tuple if there exists a sequence $\{R_H(SNR), P_e^H(SNR), R_L(SNR), P_e^L(SNR)\}$ such that $(R_H(SNR), P_e^H(SNR), R_L(SNR), P_e^L(SNR))$ is an achievable performance at each finite SNR and

$$D_H = \lim_{SNR\to\infty} -\frac{\log P_e^H(SNR)}{\log(SNR)},\ r_H = \lim_{SNR\to\infty} \frac{R_H(SNR)}{\log(SNR)} \tag{19}$$

$$D_L = \lim_{SNR\to\infty} -\frac{\log P_e^L(SNR)}{\log(SNR)},\ r_L = \lim_{SNR\to\infty} \frac{R_L(SNR)}{\log(SNR)}. \tag{20}$$

To ensure this performance we will see in Section V that a variation of the NVD criterion (6) is needed.

C. Glossary

Symbol	Interpretation
D_H	Diversity order for the message stream $\mathcal{H}$.
D_L	Diversity order for the message stream $\mathcal{L}$.
R_H	Rate for the message stream $\mathcal{H}$, for transmission from finite alphabet.
R_L	Rate for the message stream $\mathcal{L}$, for transmission from finite alphabet.
r_H	Multiplexing rate for the message stream $\mathcal{H}$ (information theoretic) $(r_H \log(SNR) = R_H)$.
r_L	Multiplexing rate for the message stream $\mathcal{L}$ (information theoretic) $(r_L \log(SNR) = R_L)$.
$X_{a,b}$	Codeword corresponding to message $\mathbf{a}$ and $\mathbf{b}$ on the streams $\mathcal{H}$ and $\mathcal{L}$ respectively.
Θ_ν	Maps a matrix of dimension $M_t \times T$ to a matrix of dimension $(\nu+1)M_t \times T$ (16).

IV. Diversity Embedded Codes: Finite Alphabet

In this section we will construct diversity embedded codes satisfying the code design criterion in Section III-B for transmission from a finite alphabet. First, some linear code constructions suitable for flat-fading channels are presented followed by a multilevel coding scheme for ISI channels which maps rank guarantees in the binary domain to rank guarantees in complex domain. We give constructions of such sets of binary matrices for transmission over an ISI channel and then specialize them for the flat-fading channel.

A. Linear Code Constructions: Flat-Fading Channels

Linear additive codeword designs are amenable to computationally-efficient lattice decoding strategies, such as the sphere decoder [9] and appear to be less susceptible to degradation in performance at low-to-moderate SNR caused by the number of nearest neighbors [7]. With this motivation, choosing the transmitted signal to be additive of the following form

$$\mathbf{X}_{\mathbf{a},\mathbf{b}} = \mathbf{X}_{\mathbf{a}} + \mathbf{X}_{\mathbf{b}}, \tag{21}$$

allows the composite code to be linear in the complex field as well. We give two code construction examples for $M_t = 4$. Constructions for 2 and 3 transmit antennas are presented in [11] and [44].

Linear Code Examples: Assuming no restrictions on which constellations $\mathcal{A}$ are used to transmit codeword $\mathbf{X}_{\mathbf{a},\mathbf{b}}$, start with a baseline code derived from the rate-$\frac{3}{4}$ orthogonal design based on Octonions [4], [41]. Therefore, the proofs of diversity order are based on symbols from a complex field and do not use properties of specific constellations.

Example 1: Here $\mathcal{H}$ comes from the message set $\{a(0), a(1), a(2)\} \in \mathcal{A}$ and $\mathcal{L}$ comes from $b(0) \in \mathcal{A}$. This implies that $R_H = \frac{3}{4}\log|\mathcal{A}|$, and $R_L = \frac{1}{4}\log|\mathcal{A}|$. The codewords can be written as

$$\mathbf{X}_{\mathbf{a},\mathbf{b}} = \mathbf{X}_{\mathbf{a}} + \mathbf{X}_{\mathbf{b}} = \begin{bmatrix} a(0) & a(1) & a(2) & b(0) \\ -a^*(1) & a^*(0) & 0 & a(2) \\ -a^*(2) & 0 & a^*(0) & -a(1) \\ 0 & -a^*(2) & a^*(1) & a(0) \end{bmatrix}. \tag{22}$$

This code can be shown to achieve full diversity of $4M_r$ for the message set $\mathcal{H}$ and diversity M_r for message set $\mathcal{L}$ (see [17] for proof details). ■

Example 2: Here $\mathcal{H}$ comes from the message set $\{a(0), a(1)\} \in \mathcal{A}$ and the message set $\mathcal{L}$ is given by $\{b(0), b(1), b(2), b(3)\} \in \mathcal{A}$. This implies that $R_H = \frac{1}{2}\log|\mathcal{A}|$, and $R_L = \log|\mathcal{A}|$. The code can be written as

$$\mathbf{X}_{\mathbf{a},\mathbf{b}} = \mathbf{X}_{\mathbf{a}} + \mathbf{X}_{\mathbf{b}} = \begin{bmatrix} a(0) & a(1) & b(2) & b(3) \\ -a^*(1) & a^*(0) & b^*(3) & -b^*(2) \\ b(0) & b(1) & a^*(0) & -a(1) \\ -b^*(1) & b(0) & a^*(1) & a(0) \end{bmatrix}. \tag{23}$$

This code can be shown to achieve diversity of $3M_r$ for message set $\mathcal{H}$ and diversity $2M_r$ for message set $\mathcal{L}$. Therefore, this example achieves the tuple, $(\frac{1}{2}\log|\mathcal{A}|, 3M_r, \log|\mathcal{A}|, 2M_r)$ (see [17] for proof details). ■

B. Multi-Level Constructions

Given an L-level binary partition of a QAM or PSK signal constellation, a space-time codeword is an array $\mathbf{K} = \{\mathbf{K}_1, \mathbf{K}_2, \ldots, \mathbf{K}_L\}$ determined by a sequence of binary matrices where matrix $\mathbf{K}_i$ specifies the space-time array at level i. A *multi-level space-time code* is defined by the choice of the constituent sets of binary matrices $\mathcal{K}_1, \mathcal{K}_2, \ldots, \mathcal{K}_L$ where $\mathbf{K}_i \in \mathcal{K}_i$. These sets of binary matrices provide rank guarantees necessary to achieve the diversity orders required for each message set.

With $|\mathcal{A}| = 2^L$, given message sets $\{\mathcal{E}_i\}_{i=1}^L$, we define the following mapping to the space-time codeword $\mathbf{X} \in \mathcal{A}^{M \times T}$

$$\{\mathcal{E}\}_{i=1}^{L} \xrightarrow{f_1} \mathbf{K} = \begin{bmatrix} K(1,1) & \dots & K(1,T) \\ \vdots & \vdots & \vdots \\ K(M,1) & \dots & K(M,T) \end{bmatrix} \xrightarrow{f_2} \mathbf{X} = \begin{bmatrix} x(1,1) & \dots & x(1,T) \\ \vdots & \vdots & \vdots \\ x(M,1) & \dots & x(M,T) \end{bmatrix} \tag{24}$$

where the matrix $\mathbf{K}$ is specified by $K(m,n) \in \{0,1\}^{\log(|\mathcal{A}|)}$ *i.e.*, a length-L binary string and $x(m,n) \in \mathcal{A}$. This construction is illustrated in Figure 3 for a constellation size of L bits. Note that the dimension of the matrix $\mathbf{X}$ or equivalently the dimension of the matrices $\mathbf{K}_i$ can be chosen to be an arbitrary integer M.

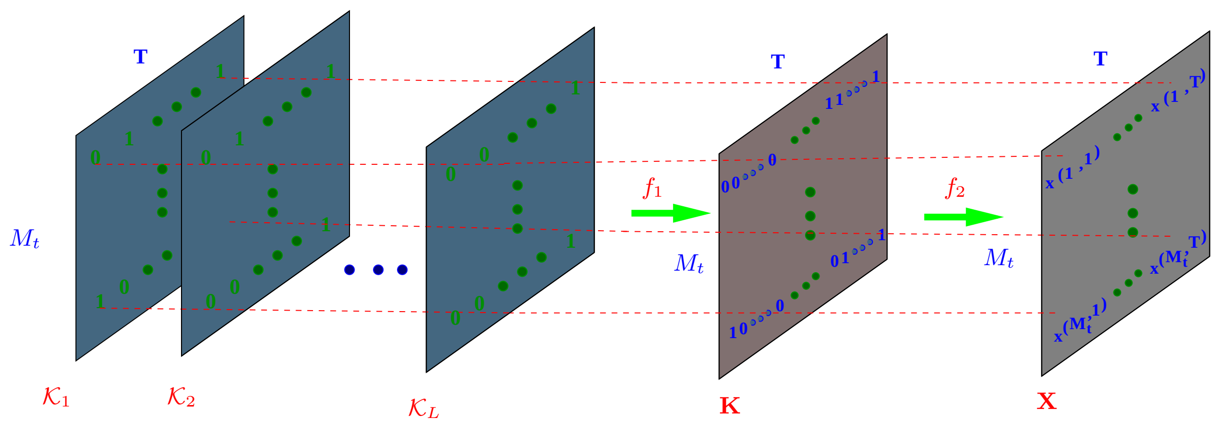


Fig. 3. Schematic representation of the multi-level code construction.

In summary, given the channel and the message set, first choose the sets of constituent matrices $\mathcal{K}_1, \mathcal{K}_2, \dots, \mathcal{K}_L$ and then the corresponding $\mathbf{K}_1, \dots, \mathbf{K}_L$. The first mapping f_1 is obtained by constructing the matrix $\mathbf{K} \in (\mathbb{F}_2^L)^{M \times T}$ whose entries are constructed by concatenating the bits from the corresponding entries in the matrices $\mathbf{K}_1, \dots .\mathbf{K}_L$ into a length-L bit-string. This is then mapped to the space-time codeword through the constellation mapper $f_2(b_0, \dots, b_{L-1}) = s$. For QAM and PSK constellations, the mapping is defined next.

- For a QAM constellation, the point s is a point in the constellation given by

$$s - c(L) \equiv \sum_{l=0}^{L-1} b_l (1-i)^l \mod (1-i)^L. \tag{25}$$

  The constant $c(L) = \frac{1}{2}$ for odd L and $\frac{1}{2}(1+i)$ for even L.

- Defining $\xi = \xi_{2^L} = \exp(2\pi i/2^L)$, points in the 2^L-PSK constellation can be represented as

$$s = \prod_{l=0}^{L-1} (\xi^{2^l})^{b_l} . \tag{26}$$

Using this sequence of L matrices, the space-time codeword can be obtained as seen in Figure 3.

Choose the sets $\mathcal{K}_l, l = 1, \ldots, L$ to have rank guarantees in the binary domain such that for any distinct pair of matrices $\mathbf{A}, \mathbf{B} \in \mathcal{K}_l$ the rank of $(\mathbf{A} - \mathbf{B})$ is at least d_l. The following theorem shows that the corresponding bits can be given rank guarantees in the complex-domain space time code.

Theorem 4.1: [17] Let $\mathcal{C}$ be a multi-level space-time code for a QAM or PSK constellation of size 2^L of dimension $M \times T$, that is determined by constituent sets of binary matrices $\mathcal{K}_l, l = 1, \ldots, L$, with binary rank guarantees $d_1 \geq d_2 \ldots \geq d_L$. Consider space-time codewords constructed from these sets using the construction in (24). This codebook satisfies the property that the minimum rank difference in the complex domain between two codewords corresponding to different input bits in the l^{th} layer is at least d_l.

Since the coding scheme in Section III-A is used, when transmitting over an ISI channel with $\nu+1$ taps, we require that the mapping in (24) satisfies the constraint that the last ν entries of the mapping lead to *given* alphabets (taken to be zero without loss of generality). Note that since the received codeword after transmission over the ISI channel can be effectively written as in (12), we need to give guarantees on the rank of $[\Theta_\nu(\mathbf{A}) - \Theta_\nu(\mathbf{B})]$ for any distinct pair of matrices $\mathbf{A}, \mathbf{B} \in \mathcal{K}_l$. Denote $\mathcal{K}_{\nu,d} \in \mathbb{F}_2^{M_t \times T}$ to be the set of binary matrices such that the last ν entries are zero and for $\mathbf{A} \neq \mathbf{B}$, the rank of $[\Theta_\nu(\mathbf{A}) - \Theta_\nu(\mathbf{B})]$ is at least $d(\nu+1)$.

An obvious way to construct $\mathcal{K}_{\nu,d}$ would be to take a single matrix in each layer which satisfies these constraints. Then choosing $\mathcal{K}_l = \mathcal{K}_{\nu,d}$, would give a rate of $R = \frac{1}{T} \log |\mathcal{K}_{\nu,d}| = 0$ on the l^{th} layer. Also from Theorem 2.2 it is known that

$$R \leq M_t - d + 1 \Rightarrow |\mathcal{K}_{\nu,d}| \leq 2^{T(M_t - d + 1)} \tag{27}$$

The following lemma shows that for some fixed $T \geq T_{thr}$ it is possible to construct binary matrices which satisfy the structure and rank guarantees with rate almost close to the maximum possible as in (27).

Lemma 4.2: [20] For block size $T \geq T_{thr} = R\nu + (M_t - 1)(\nu + 1)(2^R - 1)$, we can construct sets of binary matrices $\mathcal{K}_{\nu,d}$ such that they satisfy the following properties

- The last ν entries of all the M_t rows are equal to zero.
- For any distinct pair of matrices $\mathbf{A}, \mathbf{B} \in \mathcal{K}_{\nu,d}$ the rank of $[\Theta_\nu(\mathbf{A}) - \Theta_\nu(\mathbf{B})]$ is at least $d(\nu+1)$.
- $|\mathcal{K}_{\nu,d}| \geq 2^{T(M_t - d + 1) - \nu M_t}$.

Combining Theorem 4.1 and Lemma 4.2, we can state the formal construction guarantee for the diversity embedded code for transmission over the ISI channel as follows.

Theorem 4.3: [20] Let $\mathcal{C}$ be a multi-level space-time code for a QAM or PSK constellation of size 2^L with M_t transmit antennas that is determined by constituent sets of binary matrices $\mathcal{K}_l = \mathcal{K}_{\nu,d_l}, l = 1, \ldots, L$, such that $d_1 \geq d_2 \ldots \geq d_L$. For joint maximum-likelihood decoding, the input bits that select the codeword from the lth set $\mathcal{K}_l$ are guaranteed diversity $D_l = d_l(\nu+1)M_r$ in the complex domain when transmitted over an ISI channel with $\nu+1$ taps.

Therefore, this construction for QAM constellations achieves the rate-diversity tuple $(R_1, M_r d_1(\nu+1), \ldots, R_L, M_r d_L(\nu+1))$, with the overall equivalent single-layer code achieving rate-diversity point, $(\sum_l R_l, M_r d_L(\nu+1))$. In particular, we can construct a space-time code by choosing identical rank requirements for all the layers, *i.e.*, $d_1 = d_2 = \ldots = d_L$. Hence, the rate-diversity tradeoff for the ISI channel can be characterized as follows.

Theorem 4.4: **(Transmit-alphabet-constrained, rate-diversity tradeoff for ISI channels)** Consider transmission over a $\nu+1$ tap ISI channel with M_t transmit antennas from a QAM or PSK signal constellation $\mathcal{A}$ with $|\mathcal{A}| = 2^L$ and communication over a time period T such that $T \geq T_{thr}$. For diversity order $D_{isi} = d(\nu+1)M_r$, the rate-diversity tradeoff is given by

$$(M_t - d + 1) - \frac{\nu}{T} M_t \leq R^{eff} \leq (M_t - d + 1) \,.$$

where R_{eff} is the effective rate of transmission which includes the overhead due to the zero padding.

Note that the bounds in the above theorem are tight as $T \to \infty$. Thus, it follows that it is possible to achieve a $(\nu+1)-$fold increase in the diversity order for ISI channels for transmission over channels with multiple transmit antennas as well. The tradeoff shows that it is possible to harness all the diversity in the ISI channel with a very small rate penalty.

C. Binary Matrices

In Section IV-B we assumed the existence of sets of binary matrices $\mathcal{K}_{\nu,d}$ in Lemma 4.2. In this section we describe the construction of the sets $\mathcal{K}_{\nu,d}$ and in particular give the proof for $\mathcal{K}_{\nu,M_t}$.

Given a rate R, define the linearized polynomial

$$f(x) = \sum_{l=0}^{R-1} f_l x^{2^l}, \tag{28}$$

where $\{f_l\}_{l=0}^{R-1} \in \mathbb{F}_{2^T}$ are elements of an extension field $\mathbb{F}_{2^T}$. To develop the binary matrices define $\mathbf{c}_f \in \mathbb{F}_{2^T}^{M_t}$

$$\mathbf{c}_f = \begin{bmatrix} f(1) & f(\xi) & \ldots & f(\xi^{(M_t-1)}) \end{bmatrix}^t . \tag{29}$$

where $\xi = \alpha^{(2^R-1)(\nu+1)}$, and α is a primitive element of $\mathbb{F}_{2^T}$. Let $\mathbf{f}^{(0)}(\xi^i)$ and $\mathbf{f}^{(k)}(\xi^i)$ be the representations of $f(\xi^i)$ and $\alpha^k f(\xi^i)$ in the basis $\{\alpha^0, \alpha^1, \ldots, \alpha^{T-1}\}$, respectively *i.e.*, $\mathbf{f}^{(0)}(\xi^i), \mathbf{f}^{(k)}(\xi^i) \in \mathbb{F}_2^{1\times T}$. A matrix representation $\mathbf{C}_f \in \mathbb{F}_2^{M_t \times T}$ of $\mathbf{c}_f$ can be obtained as

$$\mathbf{C}_f = \begin{bmatrix} \mathbf{f}^{(0)t}(1) & \mathbf{f}^{(0)t}(\xi) & \ldots & \mathbf{f}^{(0)t}(\xi^{(M_t-1)}) \end{bmatrix}^t . \tag{30}$$

Note that the j^{th} row of $\mathbf{C}_f$ is given by the binary expansion of $f(\xi^{j-1}) \in \mathbb{F}_{2^T}$ in terms of the basis $\{\alpha^0, \alpha, \ldots, \alpha^{T-1}\}$. The coefficients in this basis expansion can be obtained using the trace operator described in standard textbooks on finite fields [33].

Now, in order to get the structure required in Lemma 4.2, we need to study the requirements on f so that the last ν elements in $\mathbf{C}_f$ are 0 for all the M_t rows. Using the trace [3] function, the admissible f can be represented in terms of the set $\mathcal{S}$ defined as

$$\mathcal{S} = \left\{ f : f \in \mathbb{F}_{2^T}^R, \mathrm{Tr}_{2^T/2}\left(\theta_i f(\xi^{j-1})\right) = 0 \ \forall i \in \{T-\nu, \ldots, T-1\} \text{ and } \forall j \in \{1, \ldots, M_t\} \right\}. \quad (31)$$

The cardinality of these binary sets can be shown to be lower bounded in the following Theorem.

Theorem 4.5: [20] Considering $T > (\nu+1)M_t$, a lower bound on the cardinality of the set $\mathcal{S}$ is given by $|\mathcal{S}| \geq 2^{RT-\nu M_t}$, which implies a lower bound on the effective rate of $R_{eff} = \frac{1}{T}\log|\mathcal{S}| \geq R - \frac{\nu M_t}{T}$. Theorem 4.5 implies that not much is lost, in terms of rate, by the zero padding at the end of the transmission block. Also, it can be shown that as long as $T \geq T_{thr}$, defining $\mathcal{K}_{\nu,d} = \{\mathbf{C}_f : f \in \mathcal{S}\}$, the set $\mathcal{K}_{\nu,d}$ satisfies all the properties listed in Lemma 4.2.

Simplifications for flat-fading channels: Note that for the special case of flat-fading channels, constructions of matrices exist which achieve the rate-diversity tradeoff exist for much lower value of T_{thr} [27], [34]. In particular, it is possible to construct sets of binary matrices $\mathcal{K}_{0,d}$ with $T \geq M_t$. This significantly reduces the decoding complexity and the delay while transmitting diversity embedded codes over a flat-fading channel and the construction is summarized in the following theorem.

Theorem 4.6: [27], [34] Given a rate R, define the linearized polynomial

$$f(x) = \sum_{l=0}^{R-1} f_l x^{2^l}. \quad (32)$$

where $\{f_l\}_{l=0}^{R-1} \in \mathbb{F}_{2^T}$ and α is the primitive element of $\mathbb{F}_{2^T}$. Associate the matrix $\mathbf{C}_f \in \mathbb{F}_2^{M_t \times T}$ with each such f where

$$\mathbf{C}_f = \begin{bmatrix} \mathbf{f}^{(0)t}(1) & \mathbf{f}^{(0)t}(\alpha) & \ldots & \mathbf{f}^{(0)t}(\alpha^{(M_t-1)}) \end{bmatrix}^t \quad (33)$$

and $T \geq M_t$. Defining $\mathcal{K}_{0,d} = \left\{\mathbf{C}_f : f \in \mathbb{F}_{2^T}^R\right\}$, the minimum rank distance between matrices in $\mathcal{K}_{0,d}$ is equal to $(M_t - R + 1)$.

Although the only difference between the constructions in Theorem 4.6 for flat-fading channels and the construction for ISI channels [20] is in the evaluation points of the function f, the proof techniques for the ISI case require a much more sophisticated argument. This is because of the structure and constraints

[3] Note that for any element $\beta \in \mathbb{F}_{2^T}$ the trace of the element β relative to the base field $\mathbb{F}_2$ is defined as,

$$\mathrm{Tr}_{2^T/2}(\beta) = \beta + \beta^2 + \beta^{2^2} + \ldots + \beta^{2^{T-1}}.$$

on the binary matrices imposed by the ISI channel as well as the fact that rank guarantees are required after the mapping defined in (16). Note that if space-time codes that achieve diversity order D over a flat-fading channel are used, all that can be guaranteed is that we will still achieve diversity order D over a fading ISI channel [42]. In particular [20] provides an example of a code which achieves particular points on rate-diversity tradeoff for flat-fading channels and fails to do so in the case of ISI channels. Therefore, the design of codes for fading ISI channels cannot be immediately done by using the codes for flat-fading channels.

D. Maximal-Rank Binary Codes with Toeplitz (ISI) constraints

The proof that by choosing $\mathcal{K}_{\nu,d} = \{\mathbf{C}_f : f \in \mathcal{S}\}$, the set $\mathcal{K}_{\nu,d}$ satisfies the rank properties in Lemma 4.2 for $d \neq M_t$ involves sophisticated arguments on the null space of $\Theta_\nu(\mathbf{A})$ for $\mathbf{A} \in \mathcal{K}_{\nu,d}$ (see [20]). The proof simplifies for the case of $d = M_t$, or equivalently $R = 1$, since in this case $f(x) = f_0 x$ which enables us to write $\beta f(x) = f(\beta x)$ for $\beta \in \mathbb{F}_{2^T}$. Therefore, in this section we show that that if $R = 1$, then for all $f \in \mathcal{S}, \text{rank}(\Gamma(\mathbf{C}_f)) \geq M_t(\nu+1)$ and for the proof of the general case please see [20]. In fact, for the case $R = 1$, $T_{thr} = M_t(\nu+1)$ is enough. Since $f \in \mathcal{S}$, the last ν elements in $\mathbf{C}_f$ are 0 for all the M_t rows. Therefore, $\mathbf{f}^{(k)}(\xi^i)$ is a cyclic shift by k positions of $\mathbf{f}(\xi^i)$. Hence, for $i \in \{0, 1, \dots, \nu\}$

$$\mathbf{C}_f^{(i)} = \left[\begin{array}{cccc} \mathbf{f}^{(i)t}(1) & \mathbf{f}^{(i)t}(\xi) & \dots & \mathbf{f}^{(i)t}(\xi^{(M_t-1)}) \end{array} \right]^t, \tag{34}$$

where $\mathbf{C}_f^{(i)}$ represents the matrix obtained by a cyclic shift of all the rows of the matrix $\mathbf{C}_f$ by i positions. Therefore

$$\Gamma(\mathbf{C}_f) = \left[\begin{array}{cccc} \mathbf{C}_f^t & \mathbf{C}_f^{(1)t} & \dots & \mathbf{C}_f^{(\nu)t} \end{array} \right]^t = \mathbf{U}_f \tag{35}$$

Note that the codeword matrix $\mathbf{U}_f \in \mathbb{F}_2^{(\nu+1)M_t \times T}$ can be obtained by the representation of each element of $\mathbf{u}_f$ in the basis $\{\alpha^0, \alpha^1, \dots, \alpha^{T-1}\}$ where

$$\mathbf{u}_f = \left[\begin{array}{cccc} \mathbf{c}_f^t & \alpha \mathbf{c}_f^t & \dots & \alpha^\nu \mathbf{c}_f^t \end{array} \right]^t \tag{36}$$

With this notation we can state the following theorem:

Theorem 4.7: **(Maximal rank distance codes)** Let $f(x) = f_0 x$, as in (28) with $R = 1$ and $T \geq M_t(\nu+1)$. Then for $\mathcal{S}$ defined in (31), $\forall f \in \mathcal{S}, \text{rank}(\Gamma(\mathbf{C}_f)) \geq M_t(\nu+1)$ over the binary field.

Proof: The result can be proved by contradiction. Suppose that $\mathcal{O} = \{\Gamma(\mathbf{C}_f) : f \in \mathcal{S}\}$ has rank distance less than $(\nu+1)M_t$, then there exists a vector $\mathbf{u}_f \neq \mathbf{0}$ for some $f \in \mathcal{S}$ such that the corresponding binary matrix $\Gamma(\mathbf{C}_f) = \mathbf{U}_f$ has binary rank less than $(\nu+1)M_t$ (as the code is linear). So there exists a non-trivial binary vector space $\mathcal{B} \subseteq \mathbb{F}_2^{(\nu+1)M_t}$ such that for every $\mathbf{b} \in \mathcal{B}$

$$\mathbf{b}^t \mathbf{U}_f \;=\; \mathbf{0} \iff \sum_{i=1}^{(\nu+1)M_t} b_i \mathbf{U}_f(i,j) = 0, \;\; j = 1, \dots, T, \tag{37}$$

where $\mathbf{U}_f(i,j)$ is the $(i,j)^th$ entry of $\mathbf{U}_f$ and $(\cdot)^t$ is used to denote vector transpose. Since each row of $\mathbf{U}_f$ is an expansion of the rows of $\mathbf{u}_f$ in the basis $\{\alpha^0, \alpha, \ldots, \alpha^{T-1}\}$, we can write as operations over $\mathbb{F}_{2^T}$

$$\mathbf{b}^t\mathbf{u}_f = \sum_{i=1}^{(\nu+1)M_t} b_i \mathbf{u}_f(i) = \sum_{i=1}^{(\nu+1)M_t} b_i \sum_{j=1}^{T} \mathbf{U}_f(i,j)\alpha^{j-1} = \sum_{j=1}^{T} \alpha^{j-1} \left[\sum_{i=1}^{(\nu+1)M_t} b_i \mathbf{U}_f(i,j) \right], \quad (38)$$

where the basis expansion is used. Due to the linear independence of $\{\alpha^0, \alpha, \ldots, \alpha^{T-1}\}$, it is clear from (37) and (38)that

$$\mathbf{b}^t\mathbf{U}_f = 0 \iff \mathbf{b}^t\mathbf{u}_f = 0. \quad (39)$$

Now, suppose that for $\mathbf{b} \neq \mathbf{0}$,

$$\begin{aligned} \mathbf{b}^t\mathbf{u}_f &= \sum_{i=0}^{\nu} \sum_{k=0}^{M_t-1} b_{i+k(\nu+1)} \alpha^i f(\alpha^{k(\nu+1)}) = \sum_{i=0}^{\nu} \sum_{k=0}^{M_t-1} b_{i+k(\nu+1)} \alpha^i f_0 \alpha^{k(\nu+1)} \\ &= f_0 \left(\sum_{i=0}^{\nu} \sum_{k=0}^{M_t-1} b_{i+k(\nu+1)} \alpha^{i+k(\nu+1)} \right) = 0 \end{aligned} \quad (40)$$

Thus, for every $\mathbf{b} \in \mathcal{B}$ the element $\left(\sum_{i=0}^{\nu} \sum_{k=0}^{M_t-1} b_{i+k(\nu+1)} \alpha^{i+k(\nu+1)}\right)$ is a zero of $f(x)$. Note that $\{\alpha^{i+k(\nu+1)}\}$ are linearly independent for $k \in \{0, 1, \ldots, M_t - 1\}$ and $i \in \{0, 1, \ldots, \nu\}$ as $T \geq (\nu+1)M_t$. Therefore, there is only one trivial solution to the equation (40) *i.e.*, $b_{i+k(\nu+1)} = 0$ for $i = \{0, \ldots, \nu\}, k = \{1, \ldots, \nu\}$. This contradicts the fact that the null space is non-trivial since we cannot have $\mathbf{b} \neq \mathbf{0}$ and $\mathbf{b} \in \mathcal{B}$. Hence, all matrices in $\mathcal{O}$ have rank equal to $M_t(\nu + 1)$. ■

V. Information Theoretic Embedding

In the previous section, diversity embedding codes were designed for transmission over MIMO flat-fading and ISI channels when the transmitted symbols were drawn from a finite alphabet. In this section, we consider the fundamental limits of embedding for transmission over an ISI channel when the rate of the codebook is allowed to increase with SNR. As was pointed out in Section II-B, in this framework the tradeoff between rate and reliability is captured by the D-M tradeoff. Here, we will consider only two levels of reliability but our approach can be easily generalized to any number of levels. For most of this section, we focus our attention on channels with a single degree of freedom *i.e.*, $\min(M_t, M_r) = 1$.

A. Successive Refinement

For transmission over flat-fading and ISI channels, we would like to characterize the achievable D-M tuple (r_H, D_H, r_L, D_L) as described in Section III-B. If viewed as a single-layer code, the diversity embedded code achieves rate-diversity pairs (r_H, D_H) and $(r_H + r_L, D_L)$, where it is assumed that $D_H \geq D_L$. Since it is not possible to beat the single-layer rate-diversity tradeoff, note that necessarily

$D_H \leq D^{opt}(r_H)$ and $D_L \leq D^{opt}(r_H + r_L)$ where $(r, D^{opt}(r))$ is the optimal single-layer diversity-multiplexing point corresponding to the channel. Formalizing this notion, the following definition for successive refinability can be stated:

Definition 5.1: [14] A channel is said to be successively refinable if the diversity-multiplexing tradeoff curve for transmission is successively refinable, *i.e.*, for any multiplexing gains r_H and r_L such that $r_H + r_L \leq min(M_t, M_r)$, the diversity orders

$$D_H = D^{opt}(r_H),\ D_L = D^{opt}(r_H + r_L) \tag{41}$$

are achievable, where $D^{opt}(r)$ is the optimal diversity order of the channel.

The concept of successive refinability can be visualized as in Figure V-C. For codes that are successively refinable this definition implies that one can perfectly embed a high diversity code within a high rate code.

Notation: Let $h_i^{(p,q)}$ represent the i^{th} tap coefficient between the p^{th} receive antenna and the q^{th} transmit antenna, $x^{(q)}[k]$ and $y^{(p)}[n]$ be the symbol transmitted on the q^{th} transmit antenna and the symbol received at the p^{th} receive antenna in the n^{th} time instant, respectively. Also, let $\mathbf{x}_{[a,b]}^{(q)}$ and $y_{[a,b]}^{(p)}$ be the symbols transmitted on the q^{th} transmit antenna and received at the p^{th} receive antenna over the time period a to b, *i.e.*,

$$\mathbf{y}_{[0,T-1]}^{(p)} = \begin{bmatrix} y^{(p)}[0] & y^{(p)}[1] & \ldots & y^{(p)}[T-1] \end{bmatrix}^t.$$

B. Structural Observation

Consider the same coding scheme as in Section III-A but instead of expressing the codeword as a Toeplitz matrix as in (12), the channel matrix can be expressed as a circulant matrix as follows:

$$\underbrace{\begin{bmatrix} \mathbf{y}[0] \\ \mathbf{y}[1] \\ \vdots \\ \mathbf{y}[T-\nu-1] \\ \ldots \\ \mathbf{y}[T-1] \end{bmatrix}}_{\mathbf{Y}} = \underbrace{\begin{bmatrix} \mathbf{H}_0 & \mathbf{0} & \ldots & \mathbf{0} & \mathbf{H}_\nu & \ldots & \mathbf{H}_2 & \mathbf{H}_1 \\ \mathbf{H}_1 & \mathbf{H}_0 & \ldots & \mathbf{0} & \mathbf{0} & \mathbf{H}_\nu & \ldots & \mathbf{H}_2 \\ \vdots & \vdots & & & \ldots & \ldots & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \ldots & \mathbf{0} & \mathbf{H}_\nu & \mathbf{H}_{\nu-1} & \ldots & \mathbf{H}_1 & \mathbf{H}_0 \end{bmatrix}}_{\mathbf{H}} \underbrace{\begin{bmatrix} \mathbf{x}[0] \\ \mathbf{x}[1] \\ \vdots \\ \mathbf{x}[T-\nu-1] \\ \mathbf{0} \end{bmatrix}}_{\mathbf{W}} + \underbrace{\begin{bmatrix} \mathbf{z}[0] \\ \mathbf{z}[1] \\ \vdots \\ \mathbf{z}[T-\nu-1] \\ \vdots \\ \mathbf{z}[T] \end{bmatrix}}_{\mathbf{Z}} \tag{42}$$

i.e.,

$$\mathbf{Y} = \mathbf{H}\mathbf{W} + \mathbf{Z} \tag{43}$$

where $\mathbf{Y} \in \mathbb{C}^{TM_r \times 1}$, $\mathbf{H} \in \mathbb{C}^{TM_r \times TM_t}$, $\mathbf{W} \in \mathbb{C}^{TM_t \times 1}$, $\mathbf{Z} \in \mathbb{C}^{TM_r \times 1}$. Denote $\mathbf{C} = circ\{c_1, c_2, \ldots, c_T\}$ to be the $T \times T$ circulant matrix given by

$$\mathbf{C} = \begin{bmatrix} c_1 & c_2 & c_3 & \ldots & c_{T-1} & c_T \\ c_T & c_1 & c_2 & \ldots & c_{T-2} & c_{T-1} \\ \vdots & & \vdots & \ddots & & \vdots \\ c_2 & c_3 & c_4 & \ldots & c_T & c_1 \end{bmatrix} \tag{44}$$

Rearranging this equation by permuting the rows and columns of equation (43), we get

$$\begin{bmatrix} \mathbf{y}^{(1)}_{[0,T-1]} \\ \mathbf{y}^{(2)}_{[0,T-1]} \\ \vdots \\ \mathbf{y}^{(M_r)}_{[0,T-1]} \end{bmatrix} = \underbrace{\begin{bmatrix} \mathbf{H}^{(1,1)} & \mathbf{H}^{(1,2)} & \ldots & \mathbf{H}^{(1,M_t)} \\ \mathbf{H}^{(2,1)} & \mathbf{H}^{(2,2)} & \ldots & \mathbf{H}^{(2,M_t)} \\ \vdots & \vdots & & \\ \mathbf{H}^{(M_r,1)} & \mathbf{H}^{(2,2)} & \ldots & \mathbf{H}^{(M_r,M_t)} \end{bmatrix}}_{\mathbf{H}_{circ}} \begin{bmatrix} \mathbf{x}^{(1)}_{[0,T-1]} \\ \mathbf{x}^{(2)}_{[0,T-1]} \\ \vdots \\ \mathbf{x}^{(M_t)}_{[0,T-1]} \end{bmatrix} + \mathbf{Z} \tag{45}$$

where $\mathbf{H}^{(p,q)}$ are circulant matrices given by

$$\mathbf{H}^{(p,q)} = circ\{h_0^{(p,q)}, 0, \ldots, 0, h_\nu^{(p,q)}, \ldots, h_2^{(p,q)}, h_1^{(p,q)}\}.$$

Since the $\mathbf{H}^{(p,q)}$ are circulant matrices, they can be written using the frequency-domain notation as $\mathbf{H}^{(p,q)} = \mathbf{Q}\mathbf{\Lambda}^{(p,q)}\mathbf{Q}^*$ where $\mathbf{Q}, \mathbf{Q}^* \in \mathbb{C}^{T \times T}$ are truncated DFT matrices and $\mathbf{\Lambda}^{(p,q)}$ are diagonal matrices with elements given by

$$\mathbf{\Lambda}^{(p,q)} = diag\left\{ \lambda_k^{(p,q)} : \lambda_k^{(p,q)} = \sum_{l=0}^{\nu} h_l^{(p,q)} e^{-\frac{2\pi j}{T}kl} \right\} \tag{46}$$

for $k = \{0, \ldots, (T-1)\}$. The asymptotic behaviors of these frequency-domain coefficients are related to the fading strength of the time-domain taps by the following lemma:

Lemma 5.2: [18] Consider the taps in the frequency domain in (46) given by

$$\lambda_k^{(p,q)} = \sum_{l=0}^{\nu} h_l^{(p,q)} e^{-\frac{2\pi j}{T}kl}$$

for $k = \{0, \ldots, (T-1)\}$, $p \in \{1, \ldots, M_r\}$ and $q \in \{1, \ldots, M_t\}$. For $\alpha \in (0, 1]$, define the sets $\mathcal{G}^{(p,q)}$, $\mathcal{F}^{(p,q)}(\alpha)$ and $\mathcal{M}(\alpha)$ as

$$\mathcal{G}^{(p,q)} = \{k : |\lambda_k^{(p,q)}|^2 \doteq \max_{l \in \{0,1,\ldots,\nu\}} |h_l^{(p,q)}|^2\} \tag{47}$$

$$\mathcal{F}^{(p,q)}(\alpha) = \{k : |\lambda_k^{(p,q)}|^2 \dot{\leq} SNR^{-\alpha}\}, \tag{48}$$

$$\mathcal{M}(\alpha) = \{\mathbf{h} : |h_i^{(p,q)}|^2 \dot{\leq} SNR^{-\alpha}, \forall i \in \{0, \ldots, \nu\}, \forall p \in \{1, \ldots, M_r\}, q \in \{1, \ldots, M_t\}\}. \tag{49}$$

We have the following relations on the cardinality of these sets:

(a)) Letting $\overline{\mathcal{G}^{(p,q)}}$ represent the complement of the set $\mathcal{G}^{(p,q)}$, we have $|\overline{\mathcal{G}^{(p,q)}}| \leq \nu \quad \forall p, q$. In other words, at least $T - \nu$ of the T taps in the frequency domain for each (p, q) are (asymptotically) of magnitude $\max\left(|h_0^{(p,q)}|^2, |h_1^{(p,q)}|^2, \dots, |h_\nu^{(p,q)}|^2\right)$.

(b)) Given that $\mathbf{H} \in \mathcal{M}^c(\alpha)$, for MISO channel

$$\exists p \in \{1, 2, \dots, M_t\} \ s.\ t.\ |\mathcal{F}^{(1,p)}(\alpha)| \leq \nu \tag{50}$$

and for a SIMO channel

$$\exists q \in \{1, 2, \dots, M_r\} \ s.\ t.\ |\mathcal{F}^{(q,1)}(\alpha)| \leq \nu \tag{51}$$

Here is an intuition of why such a result will hold. Consider the polynomial

$$\lambda^{(p,q)}(z) = \sum_{l=0}^{\nu} h_l^{(p,q)} z^l,$$

which evaluates to the Fourier transform for $z = e^{-\frac{2\pi j}{T}k}$. Since the polynomial of degree ν is evaluated at $z = e^{-\frac{2\pi j}{T}k}$, for $k = \{0, \dots, (T-1)\}$, at most ν values can be zero and at least $(T - \nu)$ values are bounded away from zero.

C. *Successive Refinement of ISI Tradeoff,* $min(M_t, M_r) = 1$

For transmission over the ISI channel, we propose a generic superposition coding scheme along the lines of the universal codes constructed in [24] which simplifies for special cases like the SIMO/SISO channel and flat-fading channels as shown in the Section V-D. Consider the same coding scheme as above but repeat the transmission over a period of length $(T - \nu)$ followed by ν zeros for T_b such blocks. With the same reasoning as in (45), the received symbols over the period T_bT can be written as:

$$\begin{bmatrix} \mathbf{y}^{(1)}_{[0,T-1]} & \mathbf{y}^{(1)}_{[T,2T-1]} & \cdots & \mathbf{y}^{(1)}_{[T(T_b-1),T_bT-1]} \\ \mathbf{y}^{(2)}_{[0,T-1]} & \mathbf{y}^{(2)}_{[T,2T-1]} & \cdots & \mathbf{y}^{(2)}_{[T(T_b-1),T_bT-1]} \\ \vdots & & \cdots & \vdots \\ \mathbf{y}^{(M_r)}_{[0,T-1]} & \mathbf{y}^{(M_r)}_{[T,2T-1]} & \cdots & \mathbf{y}^{(M_r)}_{[T(T_b-1),T_bT-1]} \end{bmatrix} = \mathbf{H}_{circ} \underbrace{\begin{bmatrix} \mathbf{x}^{(1)}_{[0,T-1]} & \mathbf{x}^{(1)}_{[T,2T-1]} & \cdots & \mathbf{x}^{(1)}_{[T(T_b-1),T_bT-1]} \\ \mathbf{x}^{(2)}_{[0,T-1]} & \mathbf{x}^{(2)}_{[T,2T-1]} & \cdots & \mathbf{x}^{(2)}_{[T(T_b-1),T_bT-1]} \\ \vdots & & \cdots & \vdots \\ \mathbf{x}^{(M_t)}_{[0,T-1]} & \mathbf{x}^{(M_t)}_{[T,2T-1]} & \cdots & \mathbf{x}^{(M_t)}_{[T(T_b-1),T_bT-1]} \end{bmatrix}}_{\mathbf{X}_{circ}} + \mathbf{Z} \tag{52}$$

Rearranging X_{circ} so that all the zero paddings are at the end, we will concentrate on the design of $X \in \mathbb{C}^{(T-\nu)M_t \times T_b}$ where

$$X = \begin{bmatrix} \mathbf{x}^{(1)}_{[0,T-\nu-1]} & \mathbf{x}^{(1)}_{[T,2T-\nu-1]} & \cdots & \mathbf{x}^{(1)}_{[T(T_b-1),T_bT-\nu-1]} \\ \mathbf{x}^{(2)}_{[0,T-\nu-1]} & \mathbf{x}^{(2)}_{[T,2T-\nu-1]} & \cdots & \mathbf{x}^{(2)}_{[T(T_b-1),T_bT-\nu-1]} \\ \vdots & & \cdots & \vdots \\ \mathbf{x}^{(M_t)}_{[0,T-\nu-1]} & \mathbf{x}^{(M_t)}_{[T,2T-\nu-1]} & \cdots & \mathbf{x}^{(M_t)}_{[T(T_b-1),T_bT-\nu-1]} \end{bmatrix}. \tag{53}$$

Choosing $T \geq T_{thr}$ use superposition coding such that

$$\mathbf{X} = \mathbf{X}_H + \mathbf{X}_L$$

where $\mathbf{X}_H \in \mathcal{X}_H$ and $\mathbf{X}_L \in \mathcal{X}_L$ and $\mathcal{X}_H, \mathcal{X}_L$ satisfy the following criterion:

1) For $r_H \in [0,1]$ and defining $\Delta\mathbf{X}_H = \mathbf{X}_H - \mathbf{X}_H^{'} \neq 0$, for all $\mathbf{X}_H, \mathbf{X}_H^{'} \in \mathcal{X}_H$ it is required that

$$\|\mathbf{X}_H\|_F^2 \leq (T-\nu)T_b SNR \tag{54}$$

$$\min_{\Delta\mathbf{X}_H} det\left(\Delta\mathbf{X}_H \Delta\mathbf{X}_H^*\right) \dot{\geq} SNR^{T_b-(T-\nu)r_H} \tag{55}$$

2) For $r_L + \beta \in [0,1]$, and defining $\Delta\mathbf{X}_L = \mathbf{X}_L - \mathbf{X}_L^{'} \neq 0$. for all $\mathbf{X}_L, \mathbf{X}_L^{'} \in \mathcal{X}_L$ it is required that

$$\|\mathbf{X}_L\|_F^2 \leq (T-\nu)T_b\, SNR^{1-\beta} \tag{56}$$

$$\min_{\Delta\mathbf{X}_L} det\left(\Delta\mathbf{X}_L \Delta\mathbf{X}_L^*\right) \dot{\geq} SNR^{T_b-(T-\nu)(\beta+r_L)} \tag{57}$$

Using these superposition codes for transmission, the following theorem shows that the ISI channel is almost successively refinable. The above constraints are related to the NVD criterion and from [24] it is known that codes satisfying these constraints exist.

Theorem 5.3: [15], [21] Consider a $(\nu+1)$-tap point-to-point ISI channel with a single degree of freedom *i.e.*, $\min(M_t, M_r) = 1$ and $T_b \geq T_{thr}$. The diversity-multiplexing tradeoff for this channel is successively refinable, *i.e.*, for any multiplexing gains r_H and r_L such that $r_H + r_L \leq \frac{T-\nu}{T}$ the achievable diversity orders given by $D_H(r_H)$ and $D_L(r_L)$ are bounded as

$$\begin{aligned} \max(M_t, M_r)(\nu+1)\left(1-\tilde{r}_H\right) \leq D_H(r_H) \leq \max(M_t, M_r)(\nu+1)\left(1-r_H\right), \\ \max(M_t, M_r)(\nu+1)\left(1-(\tilde{r}_H+\tilde{r}_L)\right) \leq D_L(r_L) \leq \max(M_t, M_r)(\nu+1)\left(1-(r_H+r_L)\right) \end{aligned} \tag{58}$$

where $\tilde{r}_H = \frac{T}{T-\nu} r_H$, $\tilde{r}_L = \frac{T}{T-\nu} r_L$ and T and T_b are finite and do not grow with SNR.

Hence, Theorem 5.3 shows that the best possible performance can be achieved. This implies that for ISI fading channels with a single degree of freedom, it is possible to design ideal *opportunistic* codes. The existence of (almost) ideal opportunistic codes is surprising since one would have expected the behavior for the ISI channel to be closer to the flat-fading multiple-degrees-of-freedom case, where the D-M tradeoff was not successively refinable [16]. Since ISI channels can be viewed as a set of parallel fading channels albeit with correlated fading, the correlations completely alter the characteristic of the D-M tradeoff.

The inequalities in (58) becomes tight as T is increased and therefore the D-M trade-off for the ISI channel with single degree of freedom is successively refinable as in Definition 5.1. For the flat-fading case, we can actually get successive refinement with finite block length codes as seen in [15]. An intuition of successive refinement for the flat fading channel with single degree of freedom can be obtained by observing the achievable outage plots using superposition coding in Figure V-C. We plot the achievable

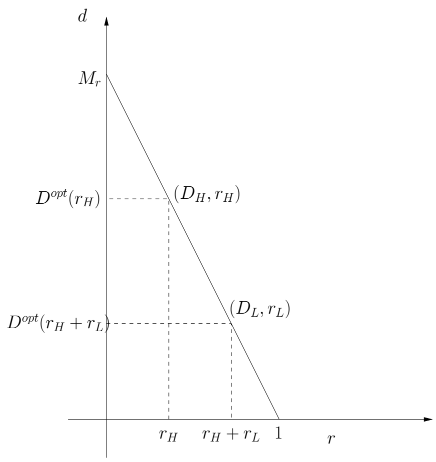


Fig. 4. Successive refinement for a flat fading channel with M_r receive antennas and one transmit antenna.

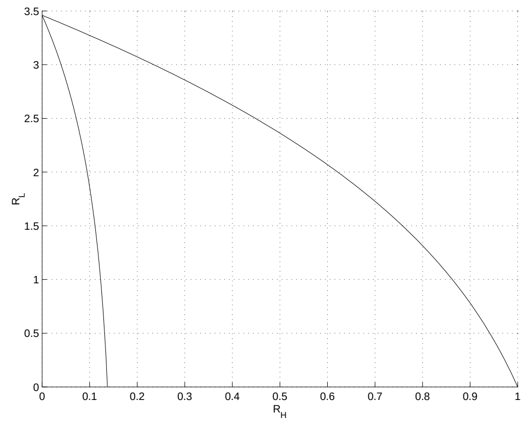


Fig. 5. Intuition for successive refinement of flat fading channels with single degree of freedom.

(R_H, R_L) points using superposition coding for different values of P_e^H and P_e^L for a fixed SNR. Note that as $\frac{P_e^H}{P_e^L}$ decreases the curve gets steeper. This implies that for a small backoff in the rate on the lower priority layer, a significant improvement is obtained in the achievable rate on the higher priority layer. This gives us the intuition that the flat fading channel with single degree of freedom is successively refinable. Note that for $D_H > D_L$ the outage probability requirements become orders of magnitude apart in the high SNR regime, and therefore diversity embedded codes are appropriate when we have such disparate reliability requirements.

D. Special Cases

In this section, we show that the coding scheme in Section V-C specializes to simpler cases for particular channels. Thus, the inequalities in (55) and (57) are sufficient but not necessary conditions for successive refinability.

- **SISO/SIMO ISI channel**: For the case of the SISO/SIMO ISI channel, assume two streams with uncoded QAM codebooks for each stream, as in [15], [18]. Let $\mathcal{X}_H$ be a QAM constellation of size $SNR^{\tilde{r}_H}$ with minimum distance $(d_{min}^H)^2 = SNR^{1-\tilde{r}_H}$. Similarly, let $\mathcal{X}_L$ be a QAM constellation of size $SNR^{\tilde{r}_L}$ with minimum distance $(d_{min}^L)^2 = SNR^{1-\beta-\tilde{r}_L}$, where $\beta > \tilde{r}_H$. Considering coding with $T_b = 1$, it is easy to see that this constellation satisfies the constraints in equations (55) and (57).
- **MISO ISI channel**: For transmission over the MISO ISI channel, we need to design codebooks satisfying Equations (55) and (57). Taking $T_b \geq TM_t$ it is possible to use sets of codebooks satisfying these properties (for example [24]), therefore establishing existence of codes with these

properties. These codes satisfy the Non Vanishing Determinant (NVD) property, which is a sufficient and necessary condition for codebooks to achieve the D-M tradeoff for flat fading channels.

The specialization for SISO/SIMO ISI channels shows that *uncoded* QAM constellations are sufficient to achieve successive refinement of D-M tradeoff.

VI. Applications

In this section, network applications of diversity embedded codes constructed in Section IV are explored and the benefits of diversity embedded codes are illustrated. In [31], the authors applied diversity-embedded codes to wireless multicasting and quantified the achievable performance gains in terms of information rate and coverage area.

A. Network utility maximization

It is important and challenging to design a network which accommodates diverse applications as different types of traffic have very different requirements for rate, reliability and delay. For example, real-time traffic needs lower delay, but non-real time traffic is delay insensitive. Therefore, traffic types with different utility functions are encountered, some of them elastic and some inelastic, all to be provisioned over the network. Diversity embedded codes allow us to allocate the packets of different traffic types to the sub-links with appropriate rate-reliability characteristics, in a way that the traffic types match the rate-reliability pairs, so that the utilities of different traffic types can be jointly optimized. The problem of how to fully utilize the different rate-reliability characteristics at the *physical layer* to support different types of traffic over a *network* and to jointly maximize their utility by extending the current framework of Network Utility Maximization (NUM) was addressed in [32].

Basic NUM assumes that each link provides a fixed-size transmission ‘pipe’ and each user’s utility is only a function of transmission rate. When the diversity-embedded codes are brought into the NUM framework, one may think of a given link as several parallel sub-links, each with different rate and reliability levels. This provides the freedom to assign user’s traffic to different sub-links so that users can have different rates, reliabilities, and delays and we refer to this scheme as capacity division. On the other hand, priority queuing puts all the packets together but annotates them according to the degree of delay sensitivity. We also modified the basic NUM framework to incorporate the delay characteristics, in addition to rate and reliability, into the utility objective function [32]. By jointly controlling the source rates, forward error correction, and ARQ, it is possible to amplify the benefits of the new codes to delay-sensitive network applications.

The question on how to allocate the physical layer resources to a mixture of traffic types, including VoIP and data traffic, with appropriate rate-reliability characteristics so that their utility are jointly maximized

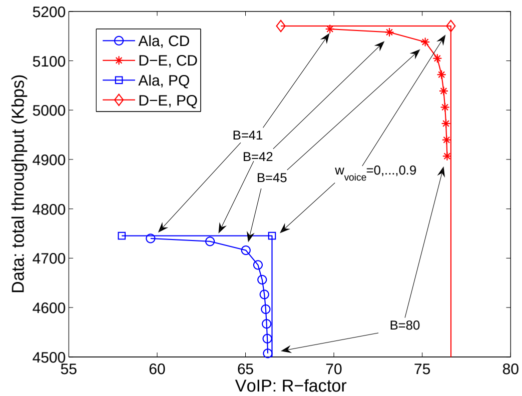


Fig. 6. Tradeoff between voice traffic and delay insensitive data traffic. Ala: Alamouti (single-diversity), D-E: Diversity-Embedded, CD: Capacity Division, PQ: Priority Queueing.

was answered in [32]. It was shown that compared with the traditional channel codes, the new diversity-embedded codes can provide higher network utilities for each traffic type *simultaneously*, with or without ARQ. This can be seen in Figure 6 where we show the tradeoff between VoIP traffic and delay insensitive data traffic. Here data traffic utility is throughput only. For capacity division, we can see the tradeoff as the capacity assigned to voice (B) varies, and for priority queueing, the tradeoff is shown as the weight of voice traffic in the objective function varies. The quality of VoIP is measured using the R-factor and the figure shows that the network with diversity-embedded codes has better tradeoff.

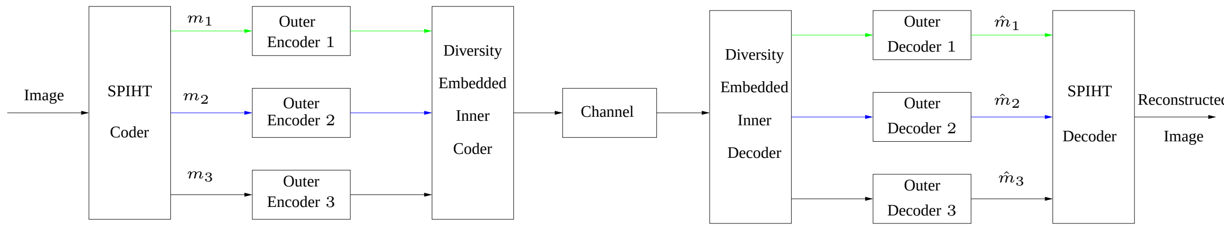
B. Wireless Image Transmission with Diversity-embedded Codes

As an application of diversity embedded codes, the UEP property designed for wireless channels is used to match it with a hierarchical source coding scheme which outputs layers of bit-stream with different priorities. The SPIHT coder [38] is one such powerful image compression algorithm that produces an embedded bit stream from which the best reconstructed images in the mean square error sense can be extracted at various bit rates. However, the SPIHT source coder is *not* robust in that the loss of a single bit in the stream renders the rest of the bits completely useless. Therefore, when integrated with the diversity embedded code, the structure needs to be designed so that this property is taken into account.

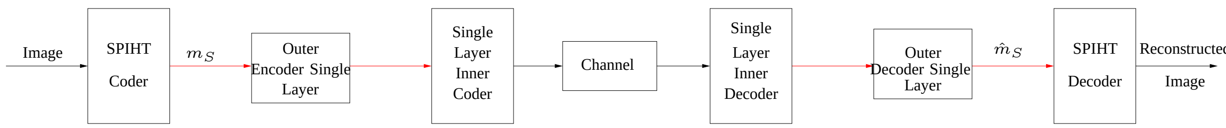
The diversity embedded code is used as an inner code along with an outer code that is designed for a particular error probability p and the overall structure is illustrated in Figure VI-B. The choice of the rate for this outer code is made by noting that given an error probability of p for transmission over a q-ary channel, the capacity is minimum when the channel is symmetric *i.e.*,

$$C_q(p) \geq C_q^{sym}(p) = \frac{1}{T}\left[\log(q) - H_b(p) - p\log(q-1)\right] \quad bpcu.$$

The rate of the outer code can be designed to be close to $C_q^{sym}(p)$ using an iterative code designed for the



(a) Concatenated diversity embedded code matched with prioritized SPIHT encoder.



(b) Concatenated single layer space-time code.

Fig. 7. Concatenated space-time codes along with SPIHT image encoder/decoder.

q-ary channel [1]. If the error probability of the inner code (*which depends on the channel realization* $\mathbf{H}$) exceeds p, an outage is declared otherwise the overall concatenated code operates reliably. This outage probability is denoted by $\mathcal{P}_{out}$.

Assume that uniform 8-PSK signaling is used for the diversity embedded code constructed using Example 1 in Section IV-A and 16-PSK Octonion codes are used for the single-layer code. Design the outer codes with rates $\mathcal{R}_1 = C_{q_1}^{sym}(0.1)$ and $\mathcal{R}_2 = C_{q_2}^{sym}(0.1)$ for the first and second layers, respectively. Similarly, for the single-layer, design the outer code with rate $\mathcal{R}_S = C_{q_S}^{sym}(0.1)$. The outage probabilities of the first layer, second layer and single layer are denoted as $\mathcal{P}_{out}^1$, $\mathcal{P}_{out}^2$ and $\mathcal{P}_{out}^S$, respectively. For the diversity embedded code, assign the most important $\mathcal{R}_1$ bits of the SPIHT encoder to layer 1 (highest diversity order layer), and correspondingly the next $\mathcal{R}_2$ bits to layer 2. In the single-layer space-time code, $\mathcal{R}_S$ bits of the SPIHT encoder are transmitted.

In Figure 8(d), the performance of a SPIHT image transmission over a slowly-fading channel when matched with a diversity embedded coder is illustrated and the quality of the received image is measured in terms of PSNR [26]. The performance of the single-layer coder is shown as well for comparison. Figure 8(d) shows the advantage of the diversity embedded code in terms of outage. Since $\max\{\mathcal{P}_1^{out}, \mathcal{P}_2^{out}\} < \mathcal{P}_S^{out}$ this means that with probability almost 0.4, there are channels for which the single-layer coder completely fails, but the diversity embedded encoder succeeds. Moreover, the design shows that even in the regime where the single-layer scheme performs better, the difference is quite small. The PSNR results can be visualized as shown in Figure 8.

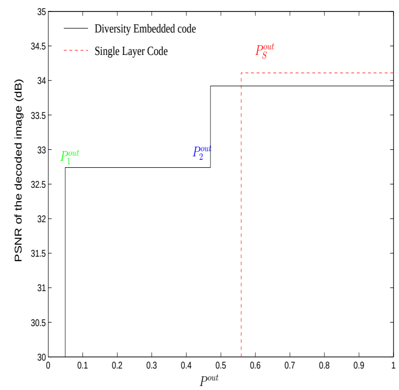
C. Rate-Delay Optimization

In this section, the delay behavior of a system integrating a rudimentary ACK/NACK feedback about the transmitted information along with space-time codes is studied. For the single-layer code the traditional

(a) Image reconstructed by decoding layer 1, with outage probability $\mathcal{P}_1^{out}$.

(b) Image reconstructed by decoding layer 1 and layer 2, with outage probability $\max\{\mathcal{P}_1^{out}, \mathcal{P}_2^{out}\}$.

(c) Image reconstructed using single layer code with outage probability $\mathcal{P}_S^{out}$.



(d) Image quality (PSNR) comparison for diversity embedded code and single layer code.

Fig. 8. Image reconstruction over a slowly fading channel.

ARQ protocol is used wherein if a packet is in error, it is re-transmitted. For the diversity embedded code, since different parts of the information receive unequal error protection, an alternative use of the ARQ can be envisaged. For two diversity levels, assume that ACK/NACK is received separately for each diversity layer. The proposed mechanism is illustrated in Figure 9. The information is sent along two streams, one on the higher-diversity level and the other on the lower-diversity level. If the packet on the higher-diversity level goes through but the lower-diversity level fails, then in the next transmission, the failed packet is sent on the higher-diversity level and therefore receives a higher "priority". Therefore, the lower-priority packet opportunistically rides along with the higher-priority packet thereby reducing the delay.

Assume a stochastic arrival of packets to be delivered at the encoder. We have an exponential inter-

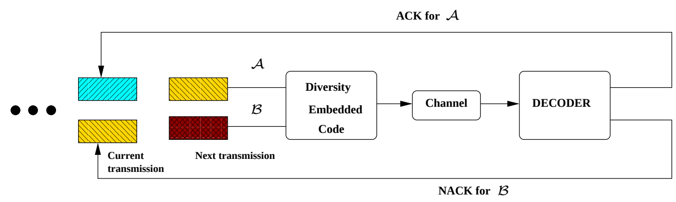


Fig. 9. ARQ mechanism for diversity embedded codes using prioritized scheduling.

arrival time point process carrying an input packet to be delivered with each arrival. The average arrival rate of the packets is set to be 95% of the average throughput of the full-diversity single-layer Octonion code. For such an arrival process, the impact of the ARQ mechanism (Figure 11) is illustrated in Figure 9 with comparison to single-layer schemes. In Figure 11, both the single-layer and the diversity embedded code are transmitted using the same transmit 8-PSK alphabet. Assuming a packet size of 243 bits is the Octonion code is used for the single-layer and the code from Example 1 in Section IV-A is used for the diversity embedded transmission.

Figure 11(b) illustrates the delay histogram for these schemes at an SNR of 19dB which shows that the delay histogram is also much better for diversity embedded codes as compared to the single-layer codes. The transmission delay for diversity embedded codes is larger, since the average time it takes to transmit one packet is larger than the single-layer codes. This disadvantage is compensated for by much better queuing delay characteristics as seen in Figure 11(d) resulting in a lower total average delay, as seen in Figure 11(a).

D. Coverage Extension

As seen in Section VI-B, it is possible to combine diversity embedded codes with an application that gives a prioritized bit stream. In this section, this functionality is examined in terms of coverage of a base-station for such applications. For simplicity, we consider a channel with only path loss and the potential range extension obtained by partial data delivery is examined. Assume a simplified path loss model where the attenuation (in dB) is given by

$$10 \log_{10} \left(\frac{P_r}{P_t} \right) = K \ (dB) - 10\gamma \log_{10} \left(\frac{d}{d_0} \right) \tag{59}$$

where K depends on the channel attenuation, d_0 is a reference distance, and γ is the path loss exponent.

For illustration, consider $M_t = 2 = T$ and 8-PSK constellation and the Alamouti code with $\mathcal{R}_S = 2.46$ for the single layer. The outer codes for the diversity-embedded code are of rates $\mathcal{R}_1 = 0.27$, $\mathcal{R}_2 = 0.51$ and $\mathcal{R}_3 = 1.39$ respectively. For a particular choice of codeword design, the received SNR (SNR_{rx}) for which the different layers can be decoded is given in the table below. Correspondingly at a received SNR of 14.4 dB, the single-layer code with information rate $\mathcal{R}_S = 2.46$ is decodable.

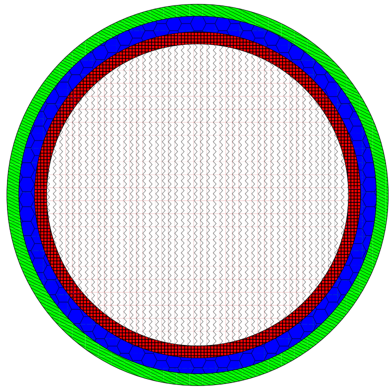


Fig. 10. Coverage increase for diversity embedded code for partial data delivery. The outermost two disks are the coverage areas for delivery of $\mathcal{R}_1, \mathcal{R}_2$ respectively. The innermost disk is the coverage area for $\mathcal{R}_3$.

$\mathcal{R}$ (bits/tx)	0.27	0.508	1.385
SNR_{rx}(dB)	12.4	13.2	15.2

These results can be translated to range increases by using the path loss model given in (59). In Figure 10, an outdoor propagation model with $\gamma = 3$ is used, for a base-station transmitting at 50W, with $d_0 = 100m$. Given these parameters, for a partial delivery of $\mathcal{R}_1 = 0.27$ bits, it is possible to get a range increase of 17% over a single-layer code operating at $\mathcal{R}_S = 2.46$ bits. The range increase for partial delivery of $\mathcal{R}_1 + \mathcal{R}_2 = 0.77$ bits over the single-layer code is 9.5%. Therefore, diversity embedded codes can improve the coverage distance of a base-station if an application can allow partial data delivery.

VII. Discussion

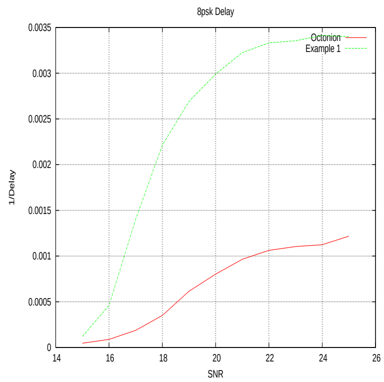
Many fundamental questions about diversity embedded codes still remain open. Characterization of optimal performance for general MIMO ISI channels, both for alphabet constrained as well as information theoretic rate growth codes is still unknown. Therefore optimal code construction for the general case remains to be found. Besides the theoretical questions, we believe that many more applications of this paradigm are yet to be discovered.

VIII. Acknowledgements

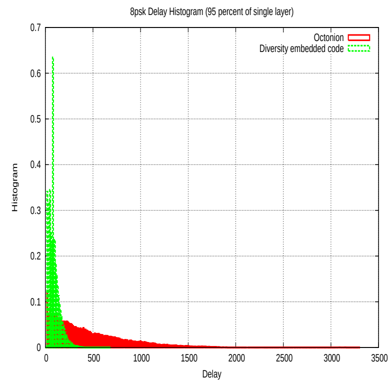
The authors would like to thank J. Chui, S. Das, Z. Islam, Y. Li, P. Rabiei, D. Tse, and D. Wang for interesting and helpful discussions about this work.

References

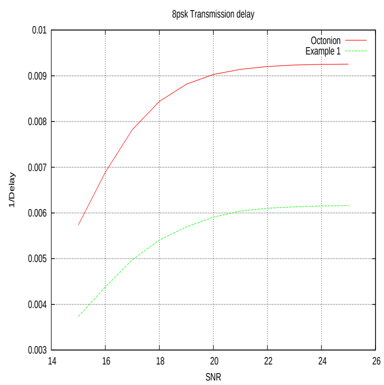
[1] A. Barg and G. Zemor, Error exponents of expander codes, *IEEE Transactions on Information Theory*, 48(6):1725–1729, June 2002.



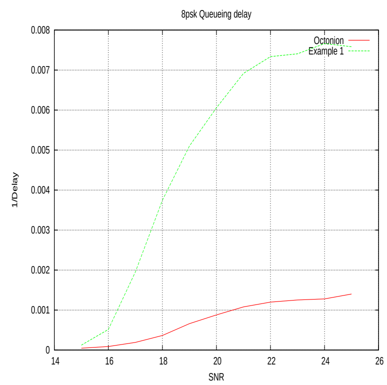
(a) Inverse average delay



(b) Delay histogram



(c) Inverse average transmission delay



(d) Inverse average queuing delay

Fig. 11. Delay comparison of diversity embedded codes with single layer codes when used with ARQ. In the figures, the transmission alphabet is fixed to be 8-PSK.

[2] E. Biglieri, J. Proakis, and S. Shamai, Fading channels: Information-Theoretic and Communications aspects, *IEEE Transactions on Information Theory*, 44(6):2619–2692, October 1998.

[3] A. R. Calderbank, Multilevel codes and multistage decoding, *IEEE Transactions on Communications*, 37(3):222 - 229, March 1989.

[4] A. R. Calderbank and A. F. Naguib, Orthogonal designs and third generation wireless communication, *Surveys in Combinatorics 2001, London Mathematical Society Lecture Note Series #288*, pages 75–107. Cambridge Univ. Press, 2001.

[5] A. R. Calderbank and N. Seshadri, Multilevel codes for Unequal Error Protection, *IEEE Transactions on Information Theory*, 39(4):1234–1248, July 1993.

[6] A. R. Calderbank, S. N. Diggavi and N. Al-Dhahir, Space-Time Signaling based on Kerdock and Delsarte-Goethals Codes, *IEEE International Conference on Communications (ICC)*, pp 483 - 487, Paris, June 2004.

[7] J. Chui and A. R. Calderbank, Effective coding gain for space-time codes, *IEEE International Symposium on Information Theory (ISIT)*, Seattle, July 2006.

[8] R. V. Cox, J. Hagenauer, N. Seshadri and C. E. W. Sundberg, Subband speech coding and matched convolutional channel coding for mobile radio channels, *IEEE Transactions on Signal Processing*, 39(8):1717–1731, August. 1991.

[9] M. O. Damen, A. Chkeif, and J. C. Belfiore, Lattice codes decoder for space-time codes, *IEEE Communication Letters*, 4:161–163, May 2000.

[10] S. Das, N. Al-Dhahir, S. Diggavi, and A. R. Calderbank, Opportunistic Space-Time Block Codes, *IEEE Vehicular Technology Conference*, pages 2025-2029, September 2005.

[11] S. Das and N. Al-Dhahir, New Diversity-Embedding STBC Constructions, *IEEE Workshop on Signal Processing Advances in Wireless Communications*, June 2006.

[12] S. N. Diggavi, N. Al-Dhahir, and A. R. Calderbank, Diversity embedding in multiple antenna communications, advances in network information theory, *DIMACS Series in Discrete Mathematics and Theoretical Computer Science*, pages 285-301, 2004.

[13] S. N. Diggavi, N. Al-Dhahir, and A. R. Calderbank, Diversity embedded space-time codes, *IEEE GLOBECOM Conference*, San Francisco, December 2003, pp 1909–1914.

[14] S. N. Diggavi and D. Tse, On successive refinement of diversity, *Allerton Conference*, October, 2004.

[15] S. N. Diggavi and D. Tse, Fundamental Limits of Diversity-Embedded Codes over Fading Channels, *IEEE International Symposium on Information Theory (ISIT)*, pp 510–514. September, 2005.

[16] S. N. Diggavi and D. Tse, On opportunistic codes and broadcast codes with degraded message sets, *IEEE Information Theory Workshop (ITW)*, pp 227–231, March, 2006.

[17] S. N. Diggavi, A. R. Calderbank, S. Dusad and N. Al-Dhahir, Diversity embedded space-time codes, *IEEE Transactions on Information Theory*, Vol 54, No. 1, pp 33-50, January 2008.

[18] S. Dusad and S. N. Diggavi, On successive refinement of diversity for fading ISI channels, *Proceedings of Allerton Conference on Communication, Control, and Computing*, Illinois, September 2006.

[19] S. Dusad, S. N. Diggavi and A. R. Calderbank, Cross layer utility of diversity embedded codes, *CISS*, Princeton, 2006.

[20] S. Dusad, S. N. Diggavi and A. R. Calderbank, Embedded Rank Distance codes for ISI channels, Submitted to *IEEE Transactions on Information Theory*, July 2007.

[21] S. Dusad and S. N. Diggavi, Successive refinement of diversity for fading ISI MISO channels, Submitted to *IEEE International Symposium on Information Theory (ISIT)*, 2008.

[22] S. Dusad and S. N. Diggavi, Successive refinement of diversity for fading ISI channels with single degree of freedom, *LICOS-REPORT-2008-008*, 2008.

[23] S. Dusad, S. N. Diggavi and A. R. Calderbank, Embedded rank distance codes for ISI channels, *LICOS-REPORT-2007-002*, May 2007.

[24] P. Elia, K. Kumar, S. Pawar, P. V. Kumar and H. F. .Lu, Explicit space-time codes achieving the diversity multplexing gain tradeoff *IEEE Transactions on Information Theory*, 52(9):3869–3884, September 2006.

[25] G. J. Foschini, Layered space-time architecture for wireless communication in a fading environment when using multi-element antennas, *Bell Labs Technical Journal*, 1(2):41–59, September 1996.

[26] J. E. Fowler, An open source software library for quantization, compression and coding, *Applications of Digital Image Processing* XXIII, Proc. SPIE 4115, pp. 294-301, August 2000. see also http://qccpack.sourceforge.net.

[27] E. Gabidulin, Theory of codes with maximum rank distance, *Probl. Per. Inform.*, 21:3–16, Jan/March 1985.

[28] H. E. Gamal, A. R. Hammons, Y. Liu, M. P. Fitz and O. Y. Takeshita, On the Design of Space-Time and Space-Frequency Codes for MIMO Frequency Selective Fading Channels, *IEEE Transactions on Information Theory*, 49(9):2277–2291, September 2003.

[29] L. Grokop and D. .N. C. Tse, Diversity Multiplexing Tradeoff in ISI channels, *IEEE International Symposium on Information Theory (ISIT)*, pp 96, Chicago, 2004.

[30] 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, *IEEE Transactions on Communications*, 47(4):527–537, April 1999.

[31] K. M. Z. Islam and N. Al-Dhahir, Hierarchical Diversity-Embedding Space-Time Block Coding, *Asilomar Conference on Signals, Systems, and Computers*, pp. 1274-1278, November 2006.

[32] Y. Li, M. Chiang, A. R. Calderbank and S. N. Diggavi, Optimal Delay-Rate-Reliability Tradeoff in Networks with Composite Links, to appear in *IEEE Transactions on Communications*, 2008.

[33] R. Lidl and H. Niederreiter, Finite Fields, *Cambridge University Press*, 1997

[34] H. F. Lu and P. V. Kumar, Rate-diversity tradeoff of space-time codes with fixed alphabet and optimal constructions for PSK modulation, *IEEE Transactions on Information Theory*, 49(10):2747–2752, October 2003.

[35] H. F. Lu and P. V. Kumar, A unified construction of space-time codes with optimal rate-diversity tradeoff, *IEEE Transactions on Information Theory*, 51(5):1709–1730, May 2005.

[36] B. Masnick and J. K. Wolf, On linear unequal error protection codes, *IEEE Transactions on Information Theory,* 13(4):600–607, October 1967.

[37] P. Rabiei and N. Al-Dhahir, Differential Diversity-Embedding Space-Time Block Coding, *Asilomar Conference on Signals, Systems, and Computers*, pp. 1291-1295, November 2006.

[38] A. Said and W. Pearlman, A new fast and efficient image codec based on set partitioning in hierarchical trees, *IEEE Transactions on Circuits and Systems for Video Technology*, (6)6, pp. 243-50, June 1996.

[39] A. Seeger, Hierarchical Channel Coding for Digital Video Broadcasting, *PhD thesis*, Technical University of Munich, Munich, Germany, 1999.

[40] I. E. Telatar, Capacity of multi-antenna Gaussian channels, *European Transactions on Telecommunications*, 10(6):585–596, November-December 1999.

[41] V. Tarokh, H. Jafarkhani, and A.R. Calderbank, Space-time block codes from orthogonal designs, *IEEE Transactions on Information Theory*, pages 1456–1467, July 1999.

[42] V. Tarokh, N. Seshadri, and A. R. Calderbank, Space-time codes for high data rate wireless communications: Performance criterion and code construction, *IEEE Transactions on Information Theory*, 44(2):744–765, March 1998.

[43] D. Tse and P. Viswanath, Fundamentals of Wireless Communication, *Cambridge University Press*, 2005.

[44] D. Wang, H. Minn, and N. Al-Dhahir, An Opportunistic STBC-OFDM Scheme with Reduced PAR in the Presence of Frequency Offset, *IEEE International Conference on Acoustics, Speech, and Signal Processing*, pp. 369-372, April 2007.

[45] L. Zheng and D. N. C. Tse, Diversity and multiplexing: A fundamental tradeoff in multiple antenna channels, *IEEE Transactions on Information Theory*, 49(5):1073–1096, May 2003.