DIMACS Series in Discrete Mathematics
and Theoretical Computer Science

Diversity Embedding in Multiple Antenna Communications

Suhas N. Diggavi, Naofal Al-Dhahir, and A. R. Calderbank

ABSTRACT. Rate and diversity impose a fundamental trade-off in space-time coding. High-rate space-time codes come at a cost of lower diversity, and high reliability (diversity) implies a lower rate. In this paper we propose a different point of view where we design high-rate space-time codes that have a high-diversity code embedded within them. This allows a form of communication where the high-rate code opportunistically takes advantage of good channel realizations whereas the embedded high-diversity code ensures that at least part of the information is received reliably. Therefore this is equivalent to multiple users transmitting information, sharing a common resource (power) and requiring different levels of diversity (error probability). We explore this point of view with design issues, along with some preliminary progress on code constructions and information-theoretic considerations.

1. Introduction

Space-time codes provide a tool to make reliable high-data rate wireless communications a reality. In [**21, 12**], it was shown that multiple transmit antennas can provide high rate communications. In [**20**], a different point of view was explored where the emphasis was on providing reliability in the form of error probability. Here, correlation of signals across the different transmit antennas was introduced to provide higher reliability albeit at the cost of transmission rate.

These two points-of-view can be combined by observing that there exists a trade-off between rate and reliability (diversity). This was explored in the context of finite-rate (fixed alphabet size) codes in [**20**] and in the context of information theoretic rate growths in [**24**]. 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 how to design transmission techniques that achieve a particular point on this trade-off. For example, the emphasis in the coding literature has been the design of maximal-diversity codes, *i.e.* one extremal point in this trade-off. In [**21, 12, 15**], the point of view taken is that of obtaining the maximal rate without consideration to reliability. There are also examples where a different point in the trade-off is achieved using codes designed for this case (see for example, [**10**]).

S N. Diggavi is the contact author. This work was presented in part in the DIMACS workshop on Network Information Theory, March 2003.

The natural question to ask is whether there exists a strategy that combines high-rate communications with high-reliability (diversity). The main idea in this paper is to propose the design of codes that achieve a high-rate, but have embedded within them a high-diversity (lower-rate) code. Clearly the overall code will still be governed by the rate-diversity trade-off, but the idea is to ensure the reliability (diversity) of at least part of the total information. A simple way to achieve this would be to time division multiplex (TDM) between codes that attain different rate-diversity points. In this paper, we explore code constructions that cannot be achieved by such simple switching strategies.

One way to view this transmission technique is from the point of view of compound channels [**8**]. Here, the information is transmitted over a channel that is unknown to the transmitter (and receiver). The strategy then is to ensure that information is decodable by the worst channel. In fact, the capacity could be strictly smaller since the transmitter needs to have a single code that is universally good for the entire class of channels [**8**]. Therefore, this is a conservative strategy, in that the transmission is robust to the worst channel, and cannot utilize the opportunity to transmit at a higher rate if a good channel is available. In [**6**], Cover proposed a broadcast strategy for transmission over compound channels. Here the idea was to superpose codewords such that the achievable rate could be strictly better than the above conservative strategy if a more benign channel is available. However, the disadvantage is that the minimum rate is smaller than that of the conservative strategy. The quasi-static fading wireless channel, where the coding is over one fading interval, is a non-ergodic channel that fits the compound channel problem. Given a random choice of channel fading values, the Shannon capacity over a class of Rayleigh fading channels is zero [**16**]. However, one can transmit at a positive rate for given outage probabilities [**16, 21, 12**]. Therefore, most transmission schemes are designed for a particular rate at a given outage probability. One exception is a broadcast strategy proposed for single-antenna fading channels which was explored in [**2, 3**], where superposition coding was used to obtain a continuum of rate and outage probabilities. In the single-antenna Rayleigh-fading case, the class of channels could be considered degraded versions of each other, depending on the received signal-to-noise ratio (SNR). The multiple-antenna channel presents a different problem where the class of channels are not degraded versions (see for example [**23**] and references therein). Moreover, it provides additional degrees of freedom that can be exploited to increase the rate and/or diversity [**21, 12, 20, 24**]. In this paper we utilize these additional degrees of freedom in order to design embedded codes. Independently, Shamai [**18**] has recently presented some extensions of the broadcast approach of single-antenna channels of [**2**] to multiple antenna channels.

The idea of unequal error protection (UEP) for information symbols has a long and rich history (see for example [**17**] and references therein). Most of the codes for UEP are designed for the Euclidean distance metric encountered in the Gaussian channel [**5**]. Another way to interpret our code-design idea is from the viewpoint of unequal error protection with the diversity metric suitable for fading channels [**20**].

The final motivation in such a design is from a networking point-of-view. Since the channel is unknown at the transmitter, such an embedding strategy makes opportunistic use of the channel. If a good channel realization is available, then we do not need significant redundancy to achieve a particular error rate target and therefore can achieve a higher information rate. This is in contrast to a strategy based on incremental redundancy codes [**14, 22**] where a low redundancy code is first sent and if an error is found then the parity symbols are sent upon feedback. In our strategy, the transmitted codebook has an embedded high-diversity code and therefore eliminates the re-transmission of parity symbols. Moreover, the rate increment is obtained without feedback about received information.

This paper is organized as follows. In Section 2, we introduce the notation used in the paper and also provide some of the background needed on space-time code design criteria as well as information-theoretic considerations. Section 3 introduces in detail the objectives of our transmission strategy and the code design criteria. We present some code constructions in Section 4 and prove some properties associated with them. The information-theoretic issues with such an embedding strategy are explored in Section 5. A numerical example in Section 6 shows the performance of one of the codes of Section 4. We conclude the paper with a brief discussion of several open issues.

2. Preliminaries

Our focus in this paper is on the quasi-static flat–fading channel where we transmit information coded over M_t transmit antennas and have M_r antennas at the receiver. Furthermore, the assumption is that the transmitter has no channel state information, whereas the receiver is able to perfectly track the channel (a common assumption, see for example [**2, 20**]). The code is designed over T transmission symbols and the received signal after demodulation and sampling can be written as

$$\mathbf{Y} = \mathbf{H}\mathbf{X} + \mathbf{Z}, \tag{2.1}$$

where $\mathbf{Y} = [\mathbf{y}(0), \ldots, \mathbf{y}(T-1)] \in \mathbb{C}^{M_r \times T}$ is the received sequence, $\mathbf{H} \in \mathbb{C}^{M_r \times M_t}$ is the quasi-static channel fading matrix, $\mathbf{X} = [\mathbf{x}(0), \ldots, \mathbf{x}(T-1)] \in \mathbb{C}^{M_t \times T}$ is the space-time coded transmission sequence with transmit power constraint P, and $\mathbf{Z} = [\mathbf{z}(0), \ldots, \mathbf{z}(T-1)] \in \mathbb{C}^{M_r \times T}$ is assumed to be additive white Gaussian noise with variance σ^2.

The coding scheme is limited to one quasi-static transmission block. Similar arguments can be made if we are allowed to code across only a *finite* number of quasi-static transmission blocks [**16**]. This allows us to view the channel in (2.1) as a non-ergodic channel since the performance is determined by single randomly chosen channel fading matrix $\mathbf{H}$.

In this context we define the notion of diversity order [**20**] as follows.

Definition 2.1. *A coding scheme which has an average error probability $\bar{P}_e(SNR)$ as a function of SNR that behaves as*

$$\lim_{SNR\to\infty} \frac{\log(\bar{P}_e(SNR))}{\log(SNR)} = -d \tag{2.2}$$

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 $\bar{P}_e(\mathrm{SNR}) \approx \mathrm{SNR}^{-d}$.

2.1. Code Design Criteria. For codes designed for a finite (and fixed) rate, one can bound the error probability by 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 [**20, 13**]. Consider a codeword sequence $\mathbf{X} = [\mathbf{x}^T(0), \dots, \mathbf{x}^T(T-1)]$ as defined in (2.1), where $\mathbf{x}(k) = [\mathbf{x}_1(k), \dots, \mathbf{x}_{M_t}(k)]^T$. The PEP between two codewords $\mathbf{x}$ and $\mathbf{e}$ can be determined by the codeword difference matrix $\mathbf{B}(\mathbf{x}, \mathbf{e})$ [**20, 13**], where

$$\mathbf{B}(\mathbf{x}, \mathbf{e}) = \begin{pmatrix} \mathbf{x}_1(0) - \mathbf{e}_1(0) & \dots & \mathbf{x}_1(T-1) - \mathbf{e}_1(T-1) \\ \vdots & \vdots & \vdots \\ \mathbf{x}_{M_t}(0) - \mathbf{e}_{M_t}(0) & \dots & \mathbf{x}_{M_t}(T-1) - \mathbf{e}_{M_t}(T-1) \end{pmatrix}.$$

For fixed rate codebook $\mathcal{C}$, the pairwise error probability between two distinct codewords $\mathbf{x}$ and $\mathbf{e}$ is given by[1] [**20**]

$$P_e(\text{SNR}, \mathbf{x} \longrightarrow \mathbf{e}) \doteq \text{SNR}^{-M_r \text{rank}(\mathbf{B}(\mathbf{x}, \mathbf{e}))}. \tag{2.3}$$

Since we are dealing with a fixed rate codebook, by using the simple union bound argument, it can be shown that the diversity order is given by [**20**]

$$d = M_r \min_{\mathbf{x} \neq \mathbf{e} \in \mathcal{C}} rank(\mathbf{B}(\mathbf{x}, \mathbf{e})) \ . \tag{2.4}$$

The error probability is determined by *both* the coding gain and the diversity order. Hence, the code design criterion prescribed in [**20**] 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. In this paper the focus is on the diversity order only, and we define corresponding criteria for embedded codebooks.

2.2. Rate vs Diversity. A natural question that arises is how many codewords can we have which allow us to attain a certain diversity order. For a flat Rayleigh fading channel, this has been examined in [**20**] where the following result was obtained.

THEOREM 2.2. *(Tarokh* et al *1998) If we use a constellation of size 2^b and the diversity order of the system is qM_r, then the rate R that can be achieved is bounded as*

$$R \leq \frac{1}{T} \log \left[A_{2^{bT}}(M_t, q) \right], \tag{2.5}$$

where $A_{2^{bT}}(M_t, q)$ is the maximum size of a code of length M_t with minimum Hamming distance q defined over an alphabet size 2^{bT}.

The discussion in Section 2.1 and Theorem 2.2 is suitable for fixed-rate (finite-alphabet) codebooks. Another point-of-view explored in [**24**], allows the rate of the codebook to increase with SNR. This viewpoint is motivated because the capacity of the multiple-antenna channel grows with SNR behaving as $\min(M_r, M_t) \log(\text{SNR})$ [**12, 24, 11**], at high SNR even for finite M_r, M_t. Therefore, [**24**] defines a multiplexing gain of a transmission scheme as follows.

---

[1] We use the notation $\doteq$ to denote exponential equality [**7**], *i.e.*, $g(\text{SNR}) \doteq \text{SNR}^a$ means that $\lim_{\text{SNR}\to\infty} \frac{\log g(\text{SNR})}{\log \text{SNR}} = a$. Moreover, if $g(\text{SNR}) \doteq f(\text{SNR})$, it means that, $\lim_{\text{SNR}\to\infty} \frac{\log g(\text{SNR})}{\log \text{SNR}} = \lim_{\text{SNR}\to\infty} \frac{\log f(\text{SNR})}{\log \text{SNR}}$.

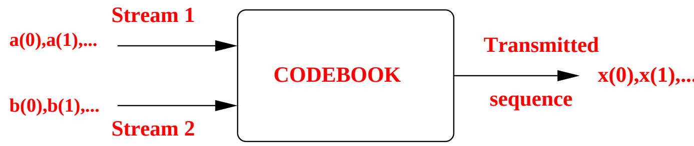

FIGURE 1. Embedded codebook

DEFINITION 2.3. *A coding scheme which has a transmission rate of $R(SNR)$ as a function of SNR is said to have a multiplexing gain r if*

$$\lim_{SNR\to\infty} \frac{R(SNR)}{\log(SNR)} = r. \tag{2.6}$$

Therefore, the system has a rate of $r\log(\text{SNR})$ at high SNR. The main result in [**24**] states that

THEOREM 2.4. *(Zheng and Tse 2002) For $T > M_t + M_r - 1$, and $K = \min(M_t, M_r)$, the optimal trade-off curve $d^*(r)$ is given by the piece-wise linear function connecting points in $(k, d^*(k)), k = 0, \dots, K$ where*

$$d^*(k) = (M_r - k)(M_t - k). \tag{2.7}$$

Therefore, both Theorems 2.2 and 2.4 show the tension between achieving high-rate and high-diversity. The question addressed in this paper is whether we can design codes with high-rate which contain within them a code which has high-diversity.

3. Diversity Embedding

The codebook structure proposed in this paper takes two information streams (see Figure 1) and outputs the transmitted sequence $\{\mathbf{x}(k)\}$. The objective is to ensure that each information stream gets the designed rate and diversity levels. In this section, we describe this approach and criteria in further detail, with specific code constructions given in Section 4.

3.1. Opportunistic Communication. The basic idea is to design a codebook which allows many levels of diversity for information streams. In particular we focus on two levels of diversity, but this procedure can be easily generalized to more levels. We design the transmitted information stream $\{\mathbf{x}(k)\}$ to depend on two information streams $\{\mathbf{a}(k)\}$ and $\{\mathbf{b}(k)\}$. Let $\mathcal{A}$ denote the message set from the first information stream and $\mathcal{B}$ denote that from the second information stream. The rates for the two message sets are, respectively, $R(\mathcal{A})$ and $R(\mathcal{B})$. The decoder jointly decodes the two message sets and we can define two error probabilities, $\bar{P}_e(\mathcal{A})$ and $\bar{P}_e(\mathcal{B})$ which denote the average error probabilities for message sets $\mathcal{A}$ and $\mathcal{B}$, respectively. We design the code such that a certain tuple (R_a, D_a, R_b, D_b) of rates and diversities are achievable, where $R_a = R(\mathcal{A}) = \frac{\log(|\mathcal{A}|)}{T}$, $R_b = R(\mathcal{B}) = \frac{\log(|\mathcal{B}|)}{T}$ and analogous to definition 2.1,

$$D_a = \lim_{\text{SNR}\to\infty} \frac{\log \bar{P}_e(\mathcal{A})}{\log(\text{SNR})}, \; D_b = \lim_{\text{SNR}\to\infty} \frac{\log \bar{P}_e(\mathcal{B})}{\log(\text{SNR})}. \tag{3.1}$$

Hence, if we examine the joint codebook, $\mathcal{C} = \{\mathcal{A}, \mathcal{B}\}$, the total rate is $R = R_a + R_b$ and the diversity is $D = \min(D_a, D_b)$. Therefore, in a baseline comparison with a single-layer codebook, the rate-diversity operating point is $(R_a + R_b, \min(D_a, D_b))$. If (R, D^*) is the optimal single-layer rate-diversity point predicted by Theorems 2.2 and 2.4, then we would like $D_a > D^*$ implying that $D_b \leq D^*$. This means that a rate $(R_a + R_b, \min(D_a, D_b))$, has a high-diversity message set $\mathcal{A}$ embedded within it. The main focus of this paper is to construct codes with this property and also to explore some characteristics of such codes. As mentioned earlier, a simple switching (TDM) strategy between codebooks designed for different rate-diversity levels could produce such a construction. However, in Section 4 we explore codes that cannot be constructed by simple switching strategies.

The reason why embedded diversity codes allow a form of opportunistic communication is that we can design the total rate $R_a + R_b$ to be large, and design D_a to be large as well. Note that there would be natural bounds on the achievable tuple (R_a, D_a, R_b, D_b), and we briefly explore this question in Section 5. This implies that we cannot get arbitrary combinations of the rate tuples. The opportunism arises because if the channel realization $\mathbf{H}$ is good, in that the channel matrix is far from singular, then the high total rate allows us to send a large amount of information through such a clear channel without knowing a-priori the channel realization at the transmitter. However, if the channel realization is not good, the embedded high-diversity message set $\mathcal{A}$ might still be decodable.

3.2. Code Design Criteria. We examine code design criteria for finite-alphabet (fixed-rate) codes. The transmitted sequence $\{\mathbf{x}(k)\}$ can be divided into codeword clusters as shown in Figure 2, where the codeword clusters shown correspond to all $\{\mathbf{x}(k)\}$ arising due to a particular message $\mathbf{a} \in \mathcal{A}$. The joint decoding of the two message sets yields candidate clusters and elements within the clusters. Given a desired tuple (R_a, D_a, R_b, D_b), we need to design message sets $\mathcal{A}, \mathcal{B}$ with size $2^{R_a T}$ and $2^{R_b T}$, respectively, to have the following properties. For message set $\mathcal{A}$, we want to design the code such that the pairwise error probability behaves as SNR^{-D_a}. Similarly for message set $\mathcal{B}$, we want to achieve diversity order of D_b. The code design criterion ensures that both these constraints are satisfied.

Let us examine a maximum-likelihood decoder jointly on the message sets $\mathcal{A}, \mathcal{B}$. Denote transmitted sequences as $\{\mathbf{x}_{\mathbf{a},\mathbf{b}}(k)\}$, with the subscripts denoting that the sequence arose due to messages $\mathbf{a} \in \mathcal{A}, \mathbf{b} \in \mathcal{B}$ as shown in Figure 1. Using (2.1) the maximum likelihood decoding metric when the channel state information is available at the receiver is given by,

$$\left(\hat{\mathbf{a}}, \hat{\mathbf{b}}\right) = \arg \min_{\mathbf{a} \in \mathcal{A}, \mathbf{b} \in \mathcal{B}} ||\mathbf{Y} - \mathbf{H}\mathbf{X}_{\mathbf{a},\mathbf{b}}||^2. \tag{3.2}$$

Therefore using (2.3) we can write the pairwise error probability on the joint message set as,

$$P_e(\mathrm{SNR}, \{\mathbf{a}, \mathbf{b}\} \longrightarrow \{\hat{\mathbf{a}}, \hat{\mathbf{b}}\}) \doteq \mathrm{SNR}^{-M_r \mathrm{rank}(\mathbf{B}(\mathbf{X}_{\mathbf{a},\mathbf{b}}, \mathbf{X}_{\hat{\mathbf{a}},\hat{\mathbf{b}}}))}. \tag{3.3}$$

The pairwise error probability between two codewords in the message set $\mathcal{A}$ is given by,

$$P_e(\mathrm{SNR}, \mathbf{a} \longrightarrow \hat{\mathbf{a}}) \doteq \mathrm{SNR}^{-M_r \min_{\mathbf{b}_1, \mathbf{b}_2 \in \mathcal{B}} \mathrm{rank}(\mathbf{B}(\mathbf{X}_{\mathbf{a},\mathbf{b}_1}, \mathbf{X}_{\hat{\mathbf{a}},\mathbf{b}_2}))}. \tag{3.4}$$

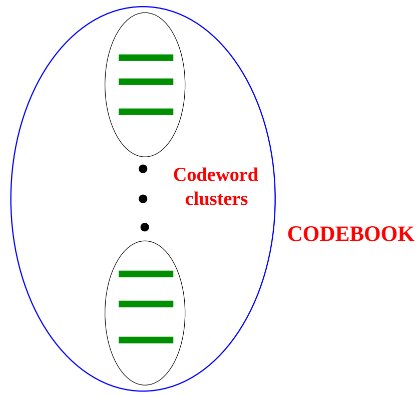

FIGURE 2. Codeword Clusters

Since we are dealing with fixed rate codebooks, by using a simple union bound argument we can write the diversity d_a for message set $\mathcal{A}$ as,

$$d_a = M_r \min_{\mathbf{a}_1 \neq \mathbf{a}_2 \in \mathcal{A}} \min_{\mathbf{b}_1, \mathbf{b}_2 \in \mathcal{B}} rank(\mathbf{B}(\mathbf{x}_{\mathbf{a}_1,\mathbf{b}_1}, \mathbf{x}_{\mathbf{a}_2,\mathbf{b}_2})) \tag{3.5}$$

Hence for the design constraint to be met, we need

$$\min_{\mathbf{a}_1 \neq \mathbf{a}_2 \in \mathcal{A}} \min_{\mathbf{b}_1, \mathbf{b}_2 \in \mathcal{B}} rank(\mathbf{B}(\mathbf{x}_{\mathbf{a}_1,\mathbf{b}_1}, \mathbf{x}_{\mathbf{a}_2,\mathbf{b}_2})) \geq D_a/M_r \tag{3.6}$$

In an identical manner, we can show for the message set $\mathcal{B}$, we need the following to hold.

$$\min_{\mathbf{b}_1 \neq \mathbf{b}_2 \in \mathcal{B}} \min_{\mathbf{a}_1, \mathbf{a}_2 \in \mathcal{A}} rank(\mathbf{B}(\mathbf{x}_{\mathbf{a}_1,\mathbf{b}_1}, \mathbf{x}_{\mathbf{a}_2,\mathbf{b}_2})) \geq D_b/M_r. \tag{3.7}$$

Basically, (3.6) implies that if we choose to transmit a particular message $\mathbf{a} \in \mathcal{A}$, regardless of which message is chosen in message set $\mathcal{B}$, we are ensured a diversity level of D_a for this message set. A similar argument will hold for message set $\mathcal{B}$.

Note that this criterion focuses *only* on the diversity level achieved by the code design. In general, the error probability depends *both* on the diversity order defined in Definition 2.1 as well as a coding gain [**20**]. The coding gain could be an important issue especially in low to moderate SNR regimes. Also, the code design criterion is suitable only for fixed rate codebooks. When one needs to design codebooks whose rates grow with SNR [**24**], alternative code design criteria could be formulated.

3.3. Challenges. There are several challenges in designing codes that meet the criteria prescribed in Section 3.2. The first obvious challenge is to design codes that can guarantee the desired tuple (R_a, D_a, R_b, D_b). The second challenge arises due to decoding complexity. In principle, we need to jointly decode both message sets $\mathcal{A}, \mathcal{B}$, and since the design is to have a high total rate $(R_a + R_b)$, the joint decoding could be quite expensive. Finally, the question arises on whether the designs are extremal, in that they achieve a point on the boundary of the achievable tuples. This requires the characterization of the achievable tuples in a manner similar to Theorems 2.2 and 2.4 which was done for the case of a single layer of transmission.

There are several natural approaches that one can attempt. The first guess would be a multilevel transmission technique similar to that done for Gaussian channels [**5**]. In this case, the idea would be based on set-partitioning multi-dimensional signal sets which are *not* necessarily generated as a product of partitions from lower-dimensional signal sets. This can be accomplished as follows. In order to illustrate the partitioning, we take the example of 4PSK signal sets. Given that M_t 4PSK signals are to be transmitted, we need to partition the $M_t \times$ 4PSK signal into a partition chain. Therefore, there are 4^{M_t} signal points in the original signal set. We partition this into $2M_t$ levels, with the l^{th} level containing 2^l disjoint subsets, where we have denoted the original signal set with level $l = 0$. We choose a partition chain, $S_0|S_1|\ldots|S_r$ where $r \leq 2M_t$. Then, codes $\{\mathcal{C}_i\}_{i=1}^r$ are used to choose elements from the partition chain. The advantage is that multi-stage decoding could be employed to reduce the decoding complexity. The main disadvantage is that like all trellis codes, the complexity is still quite large.

In [**10**] and references therein, several approaches based on constellation rotation are used to guarantee diversity. These methods are based on codes that are linear over the complex field and have the advantage of being decodable using a sphere decoder [**9**] which has an average complexity that could be polynomial in the rate and therefore is attractive for decoding high-rate codes. However, in all approaches using rotated codes, there is a significant increase in the transmitted alphabet size (constellation expansion) which we attempt to avoid. But the restriction to codes that are linear in the complex field is attractive and we study such designs in Section 4. Our designs do *not* expand the transmitted alphabet size.

4. Code Constructions

In this Section we restrict our attention to code designs that are linear over the complex field in order to be able to decode them efficiently using a sphere decoder [**9**]. In Section 4.1 we describe some general properties of the code constructions. Some specific constructions along with their properties are given in Section 4.2.

4.1. Additive Codewords Model. We examine code designs that are additive, in that the transmitted signal is such that

$$\mathbf{X}_{\mathbf{a},\mathbf{b}} = \mathbf{X}_{\mathbf{a}} + \mathbf{X}_{\mathbf{b}}. \tag{4.1}$$

This allows the composite code to be linear in the complex field as well. Given the design criterion in (3.5), it is clear that to ensure that the message set $\mathcal{A}$ gets a diversity level of D_a, a necessary (but *not* sufficient) condition is that

$$\min_{\mathbf{a}_1 \neq \mathbf{a}_2 \in \mathcal{A}} rank(\mathbf{X}_{\mathbf{a}_1} - \mathbf{X}_{\mathbf{a}_2}) \geq D_a/M_r, \tag{4.2}$$

due to the linearity of the code in (4.1). The code design should be such that any perturbation of $\mathbf{x}_a$ due to message set $\mathcal{B}$ does not disturb this rank property.

We can now apply the code design criterion developed in Section 3.2 to the additive codewords model. For illustration, consider the case where we want $D_a = M_r M_t$, *i.e.* maximal diversity. Then we clearly need $(\mathbf{X}_{\mathbf{a}_1} - \mathbf{X}_{\mathbf{a}_2})$ to have full rank for $\mathbf{a}_1 \neq \mathbf{a}_2$. We also need to ensure that $(\mathbf{X}_{\mathbf{a}_1} - \mathbf{X}_{\mathbf{a}_2}) + (\mathbf{X}_{\mathbf{b}_1} - \mathbf{X}_{\mathbf{b}_2})$ is still full rank for *any* $\mathbf{b}_1, \mathbf{b}_2 \in \mathcal{B}$. This translates to the condition

$$\nexists \mathbf{v} \in \mathbb{C}^{1\times M_t} \text{ such that } \mathbf{v}(\mathbf{X}_{\mathbf{b}_1} - \mathbf{X}_{\mathbf{b}_2})(\mathbf{X}_{\mathbf{a}_1} - \mathbf{X}_{\mathbf{a}_2})^{\dagger} = -\mathbf{v}, \quad \forall \mathbf{b}_1, \mathbf{b}_2 \in \mathcal{B}, \mathbf{a}_1 \neq \mathbf{a}_2 \in \mathcal{A} \tag{4.3}$$

where[2] $(\mathbf{Q})^{\dagger} = \mathbf{Q}^*(\mathbf{Q}\mathbf{Q}^*)^{-1}$ denotes the pseudo-inverse of a matrix. In cases of a square design, this translates to a question about existence of certain eigenvalues of the matrix defined in (4.3). Also, for square designs, the condition in (4.3) can equivalently be stated in terms of the right eigenvector in $\mathbb{C}^{M_t}$. The condition in (4.3) allows us to eliminate designs that perturb the full rank design for message set $\mathcal{A}$.

In general, the additive codeword model is amenable to computationally simpler lattice decoding strategies, such as the sphere decoder [**9**]. Moreover, rank conditions may be easier to verify under these constraints. These reasons motivate us to examine constructions based on this model.

4.2. Code Examples. In this section we provide four code constructions that illustrate the design trade-off involved in diversity-embedded codes. We restrict our attention to examples with $M_t = 4$. We also start with a baseline code derived from the rate-$\frac{3}{4}$ orthogonal design based on Octonions [**19, 4**]. In all these examples, there are no restrictions on which constellations $\mathcal{S}$ are used to transmit codeword $\mathbf{X}_{\mathbf{a},\mathbf{b}}$. 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 $\mathcal{A}$ comes from the message set $\{a(0), a(1), a(2)\} \in \mathcal{S}$ and $\mathcal{B}$ comes from $b(0) \in \mathcal{S}$. This implies that $R_a = \frac{3}{4}\log|\mathcal{S}|$, and $R_b = \frac{1}{4}\log|\mathcal{S}|$.

$$\mathbf{X}_{\mathbf{a},\mathbf{b}} = \mathbf{X}_{\mathbf{a}} + \mathbf{X}_{\mathbf{b}} = \begin{bmatrix} a(0) & a(1) & a(2) & b(0) \\ -a^*(1) & a^*(0) & 0 & a(2) \\ -a^*(2) & 0 & a^*(0) & -a(1) \\ 0 & -a^*(2) & a^*(1) & a(0) \end{bmatrix}. \tag{4.4}$$

This code can be shown to achieve full diversity of $4M_r$ for the message set $\mathcal{A}$ and diversity M_r for message set $\mathcal{B}$. The proof of this is simple and we illustrate it in order to demonstrate how the criterion in (3.5) can be applied. Let us denote $\tilde{\mathbf{a}} = \mathbf{a}_1 - \mathbf{a}_2$ and $\tilde{\mathbf{b}} = \mathbf{b}_1 - \mathbf{b}_2$. Since this is a square design, and $\mathbf{X}_{\mathbf{a}}$ is a unitary matrix, the condition given in (4.3) translates to

$$\nexists \mathbf{v} \text{ such that for } \tilde{\mathbf{a}} \neq \mathbf{0} \quad \mathbf{X}_{\tilde{\mathbf{b}}}\mathbf{X}^*_{\tilde{\mathbf{a}}}\mathbf{v} = -||\tilde{\mathbf{a}}||^2\mathbf{v}, \tag{4.5}$$

for all $\tilde{\mathbf{b}}$, where $||\tilde{\mathbf{a}}||^2 = |a(0)|^2 + |a(1)|^2 + |a(2)|^2 \neq 0$. The condition in (4.5) means that $\mathbf{X}_{\tilde{\mathbf{b}}}\mathbf{X}^*_{\tilde{\mathbf{a}}}$ should not have a negative real eigenvalue. It can easily be shown that for this design all the eigenvalues of this matrix are zero and hence the condition is met. At first glance, this design might seem uninteresting since we know that according to Theorem 2.2 (see [**20**]) there exist codes with rate $\log|\mathcal{S}|$ and diversity $M_r M_t$. However, we do not know of codes linear in the complex field, which are not designed for specific signal constellations, and have *no* constellation expansion, with such a property.

■

[2] For a matrix $\mathbf{X}$, we use the notation X^* to denote the Hermetian transpose of the matrix, which for scalars is just complex conjugation.

Example 2: Here $\mathcal{A}$ comes from the message set $\{a(0), a(1)\} \in \mathcal{S}$ and $\mathcal{B}$ comes from $\{b(0), b(1), b(2), b(3)\} \in \mathcal{S}$. This implies that $R_a = \frac{1}{2}\log|\mathcal{S}|$, and $R_b = \log|\mathcal{S}|$, leading to a total rate of $R_a + R_b = \frac{3}{2}\log|\mathcal{S}|$.

$$\mathbf{X}_{\mathbf{a},\mathbf{b}} = \mathbf{X}_{\mathbf{a}} + \mathbf{X}_{\mathbf{b}} = \begin{bmatrix} a(0) & a(1) & 0 & 0 \\ -a^*(1) & a^*(0) & 0 & 0 \\ b(0) & b(1) & a^*(0) & -a(1) \\ b(2) & b(3) & a^*(1) & a(0) \end{bmatrix}. \tag{4.6}$$

This code can be shown to achieve full diversity of $4M_r$ for the message set $\mathcal{A}$ and diversity M_r for message set $\mathcal{B}$. Therefore, this example achieves the tuple, $(\frac{1}{2}\log|\mathcal{S}|, 4M_r, \log|\mathcal{S}|, M_r)$. The proof of this claim is quite simple and follows a similar argument as in Example 1. In this example the total rate is $\frac{3}{2}\log|\mathcal{S}|$, with a composite diversity of M_r, however it has a code of maximal diversity and rate $\frac{1}{2}\log|\mathcal{S}|$ embedded within it. Another question is whether such a code can be constructed by a TDM strategy. This is not possible for the following reason. A full-diversity code needs to occupy at least M_t symbols, and hence for a square design $T = M_t$. This implies that we cannot construct a switching scheme that has a full diversity code and has a rate beyond $\log|\mathcal{S}|$.

■

Example 3: Here $\mathcal{A}$ comes from the message set $\{a(0), a(1)\} \in \mathcal{S}$ and $\mathcal{B}$ comes from $\{b(0), b(1), b(2), b(3)\} \in \mathcal{S}$. This implies that $R_a = \frac{1}{2}\log|\mathcal{S}|$, and $R_b = \log|\mathcal{S}|$, identical to Example 2.

$$\mathbf{X}_{\mathbf{a},\mathbf{b}} = \mathbf{X}_{\mathbf{a}} + \mathbf{X}_{\mathbf{b}} = \begin{bmatrix} a(0) & a(1) & b(2) & b(3) \\ -a^*(1) & a^*(0) & b^*(3) & -b^*(2) \\ b(0) & b(1) & a^*(0) & -a(1) \\ -b^*(1) & b(0) & a^*(1) & a(0) \end{bmatrix}. \tag{4.7}$$

This code can be shown to achieve diversity of $3M_r$ for the message set $\mathcal{A}$ and diversity $2M_r$ for message set $\mathcal{B}$. Therefore, this example achieves the tuple, $(\frac{1}{2}\log|\mathcal{S}|, 3M_r, \log|\mathcal{S}|, 2M_r)$. The proof of this involves a set closely related to the quaternions (used in the Alamouti code [**1**]). For $a(0), a(1) \in \mathbb{C}$, define the following sets

$$\mathcal{Q} = \begin{bmatrix} a(0) & a(1) \\ -a^*(1) & a^*(0) \end{bmatrix}, \ \tilde{\mathcal{Q}} = \begin{bmatrix} a(0) & a(1) \\ a^*(1) & -a^*(0) \end{bmatrix}. \tag{4.8}$$

The set $\mathcal{Q}$ is a multiplicative group (called the quaternionic group [**4**]), and $\tilde{\mathcal{Q}}$ is *not* a multiplicative group. However, it has the following properties

$$\begin{aligned} \mathbf{F} \in \tilde{\mathcal{Q}}, &\quad \mathbf{F}^{-1} \in \tilde{\mathcal{Q}} \\ \mathbf{E}, \mathbf{F} \in \tilde{\mathcal{Q}}, &\quad \mathbf{EF} \in \mathcal{Q}, \\ \mathbf{E} \in \mathcal{Q}, \mathbf{F} \in \tilde{\mathcal{Q}}, &\quad \mathbf{EF} \in \tilde{\mathcal{Q}} \\ \mathbf{E} \in \tilde{\mathcal{Q}}, \mathbf{F} \in \mathcal{Q}, &\quad \mathbf{EF} \in \tilde{\mathcal{Q}}. \end{aligned} \tag{4.9}$$

Let us index elements of the sets $\mathbf{E} \in \mathcal{Q}$ and $\mathbf{F} \in \tilde{\mathcal{Q}}$ in terms of the indeterminates $\{u, v\}$ as $\mathbf{E}[u, v]$ and $\mathbf{F}[u, v]$. Using this notation, we can rewrite the code in (4.7) as,

$$\mathbf{X}_{\mathbf{a},\mathbf{b}} = \mathbf{X}_{\mathbf{a}} + \mathbf{X}_{\mathbf{b}} = \begin{bmatrix} \mathbf{E}[a(0), a(1)] & \mathbf{F}[b(2), b(3)] \\ \mathbf{E}[b(0), b(1)] & \mathbf{E}[a^*(0), -a(1)] \end{bmatrix}, \tag{4.10}$$

where $\mathbf{E} \in \mathcal{Q}$ and $\mathbf{F} \in \tilde{\mathcal{Q}}$. Given two distinct candidate message sets $\{\mathbf{a}_1, \mathbf{b}_1\}$ and $\{\mathbf{a}_2, \mathbf{b}_2\}$, let us denote $\tilde{\mathbf{a}} = \mathbf{a}_1 - \mathbf{a}_2$ and $\tilde{\mathbf{b}} = \mathbf{b}_1 - \mathbf{b}_2$. To show that $D_a = 3M_r$, we need to show that if $\tilde{\mathbf{a}} \neq \mathbf{0}$, then $\min_{\tilde{\mathbf{a}} \neq \mathbf{0}} \min_{\tilde{\mathbf{b}}} \text{rank}(\mathbf{X}_{\tilde{\mathbf{a}},\tilde{\mathbf{b}}}) = 3$. By choosing $\tilde{b}(0) = \tilde{a}(0), \tilde{b}(1) = \tilde{a}(1)$ and $\tilde{b}(2) = \tilde{a}^*(0), \tilde{b}(3) = -\tilde{a}(1)$, we get $\det(\mathbf{X}_{\tilde{\mathbf{a}},\tilde{\mathbf{b}}}) = 0$, and hence the diversity for message set $\mathcal{A}$ is at most $3M_r$, *i.e.* $D_a \leq 3M_r$. But, the matrix $\begin{bmatrix} \mathbf{E}[\tilde{a}(0), \tilde{a}(1)] \\ \mathbf{E}[\tilde{b}(0), \tilde{b}(1)] \end{bmatrix}$ for $\mathbf{E} \in \mathcal{Q}$, has orthogonal columns and therefore has full column rank (*i.e.* rank of 2). This shows that $D_a \geq 2M_r$. We show that $D_a = 3M_r$ by contradiction. Let us assume that $D_a = 2M_r$, *i.e.*, there exists $\tilde{\mathbf{a}} \neq \mathbf{0}, \tilde{\mathbf{b}}$ such that $\text{rank}(\mathbf{X}_{\tilde{\mathbf{a}},\tilde{\mathbf{b}}}) = 2$. When either $\mathbf{E}[\tilde{b}(0), \tilde{b}(1)] = \mathbf{0}$ or when $\mathbf{F}[\tilde{b}(2), \tilde{b}(3)] = \mathbf{0}$, one can easily see that this cannot happen. Suppose it is true for $\mathbf{E}[\tilde{b}(0), \tilde{b}(1)] \neq \mathbf{0}$ and $\mathbf{F}[\tilde{b}(2), \tilde{b}(3)] \neq \mathbf{0}$. This implies that there exists $\Gamma \in \mathbb{C}^{2\times 2}$ such that,

$$\begin{bmatrix} \mathbf{E}[\tilde{a}(0), \tilde{a}(1)] \\ \mathbf{E}[\tilde{b}(0), \tilde{b}(1)] \end{bmatrix} = \begin{bmatrix} \mathbf{F}[\tilde{b}(2), \tilde{b}(3)] \\ \mathbf{E}^*[\tilde{a}^*(0), -\tilde{a}(1)] \end{bmatrix} \Gamma. \tag{4.11}$$

Hence this implies that

$$\mathbf{E}[\tilde{b}(0), \tilde{b}(1)] = \mathbf{E}^*[\tilde{a}^*(0), -\tilde{a}(1)]\mathbf{F}^{-1}[\tilde{b}(2), \tilde{b}(3)]\mathbf{E}[\tilde{a}(0), \tilde{a}(1)],$$

which cannot happen due to properties in (4.9) resulting in the contradiction. Hence this shows that $D_a = 3M_r$. Due to the orthogonality of columns of $\begin{bmatrix} \mathbf{E}[\tilde{a}(0), \tilde{a}(1)] \\ \mathbf{E}[\tilde{b}(0), \tilde{b}(1)] \end{bmatrix}$ and those in $\begin{bmatrix} \mathbf{F}[\tilde{b}(2), \tilde{b}(3)] \\ \mathbf{E}^*[\tilde{a}^*(0), -\tilde{a}(1)] \end{bmatrix}$ it is clear that $D_b = 2M_r$.

The contrast with Example 2 is that although the rates for each layer are identical, here we achieve a higher diversity for the message set $\mathcal{B}$ but at the cost of a lower diversity for the message set $\mathcal{A}$. In this example as well, we cannot construct a TDM code that achieves this. If we need a code of diversity 3, it needs to occupy at least 3 time symbols. Hence for block size $T = M_t = 4$, we are left with 1 symbol in which to obtain diversity two and rate $\log|\mathcal{S}|$, which is clearly not possible.

■

The next example illustrates the case of a non-square design.

Example 4: In this example, $M_t = 4$ and the block size is $T = 5$ symbols, giving rise to a non-square code matrix design. Here $\mathcal{A}$ comes from the message set $\{a(0), a(1)\} \in \mathcal{S}$ and $\mathcal{B}$ comes from $\{b(0), b(1), b(2), b(3)\} \in \mathcal{S}$. This implies that $R_a = \frac{2}{5}\log|\mathcal{S}|$, and $R_b = \frac{4}{5}\log|\mathcal{S}|$, leading to a total rate of $R_a + R_b = \frac{6}{5}\log|\mathcal{S}|$.

$$\mathbf{X}_{\mathbf{a},\mathbf{b}} = \mathbf{X}_{\mathbf{a}} + \mathbf{X}_{\mathbf{b}} = \begin{bmatrix} a(0) & -a^*(1) & b(0) & b(2) & -b^*(3) \\ a(1) & a^*(0) & b(1) & b(3) & b^*(2) \\ 0 & 0 & a^*(0) & a^*(1) & -b^*(1) \\ 0 & 0 & -a(1) & a(0) & b^*(0) \end{bmatrix}. \tag{4.12}$$

This code can be shown to achieve diversity of $4M_r$ for the message set $\mathcal{A}$ and diversity $2M_r$ for message set $\mathcal{B}$. Therefore, this example achieves the tuple, $(\frac{2}{5}\log|\mathcal{S}|, 4M_r, \frac{4}{5}\log|\mathcal{S}|, 2M_r)$. Given the construction from Example 2, we can easily prove that for $\tilde{\mathbf{a}} \neq \mathbf{0}$, $\text{rank}(\mathbf{X}_{\tilde{\mathbf{a}},\tilde{\mathbf{b}}}) = 4$, and hence for message set $\mathcal{A}$ we get a diversity order of $4M_r$. To prove the claimed diversity order for message set $\mathcal{B}$, we need to show the following

$$\min_{\substack{\tilde{\mathbf{a}} \\ \tilde{\mathbf{b}} \neq \mathbf{0}}} \operatorname{rank}(\mathbf{X}_{\tilde{\mathbf{a}},\tilde{\mathbf{b}}}) = 2. \tag{4.13}$$

Clearly, for $\tilde{a}(0) = \tilde{a}(1) = \tilde{b}(0) = \tilde{b}(1) = 0$, $\operatorname{rank}(\mathbf{X}_{\tilde{\mathbf{a}},\tilde{\mathbf{b}}}) = 2$ and hence the minimal rank is at most 2. We prove the result by contradiction. Suppose there exists $\tilde{\mathbf{a}}$ and $\tilde{\mathbf{b}} \neq \mathbf{0}$ such that $\operatorname{rank}(\mathbf{X}_{\tilde{\mathbf{a}},\tilde{\mathbf{b}}}) = 1$. This implies that for such a choice, all the rows are multiples of one another, *i.e.*,

$$\begin{bmatrix} \tilde{a}(0) \\ -\tilde{a}^*(1) \\ \tilde{b}(0) \\ \tilde{b}(2) \\ -\tilde{b}^*(3) \end{bmatrix} = \alpha \begin{bmatrix} \tilde{a}(1) \\ \tilde{a}^*(0) \\ \tilde{b}(1) \\ \tilde{b}(3) \\ \tilde{b}^*(2) \end{bmatrix}, \quad \begin{bmatrix} 0 \\ 0 \\ \tilde{a}^*(0) \\ \tilde{a}^*(1) \\ -\tilde{b}^*(1) \end{bmatrix} = \beta \begin{bmatrix} 0 \\ 0 \\ -\tilde{a}(1) \\ \tilde{a}(0) \\ \tilde{b}^*(0) \end{bmatrix}, \quad \begin{bmatrix} \tilde{a}(1) \\ \tilde{a}^*(0) \\ \tilde{b}(1) \\ \tilde{b}(3) \\ \tilde{b}^*(2) \end{bmatrix} = \gamma \begin{bmatrix} 0 \\ 0 \\ \tilde{a}^*(0) \\ \tilde{a}^*(1) \\ -\tilde{b}^*(1) \end{bmatrix}$$

For the above to be true, we would need $\tilde{a}(0) = 0 = \tilde{a}(1)$ and $\tilde{b}(0) = \tilde{b}(1) = \tilde{b}(2) = \tilde{b}(3) = 0$, which results in a contradiction. Hence we get the result claimed in (4.13), giving $D_b = 2M_r$.

This code also cannot be constructed using a switching strategy. A full-diversity code needs to occupy at least M_t symbols, and hence the message set $\mathcal{A}$ needs to occupy $M_t = 4$ symbols. Hence for block size $T = 5$, we are left with 1 symbol in which to obtain diversity two and rate $\frac{4}{5}\log|\mathcal{S}|$, which is clearly not possible. ■

5. Achievable rates

In this section we study a characterization of the achievable tuples (R_a, D_a, R_b, D_b). There can be two types of characterizations, one based on fixed-rate (finite-alphabet) codes which resemble the characterization of Theorem 2.2 [**20**], and the other a rate-growth characterization akin to Theorem 2.4 [**24**]. We present a characterization based on rate-growth in this section with the caveat that the rate growths investigated here are different from the fixed rate values given code constructions of Section 4.

5.1. Characterization. In the rate growth characterization, identical to Definition 2.3 we can define two multiplexing gains r_a, r_b for message sets $\mathcal{A}$ and $\mathcal{B}$, respectively. Therefore, the tuple in this case would be (r_a, D_a, r_b, D_b) which would be a different characterization from the tuple examined in a fixed-rate case. We also assume that the codeword length is large enough to make average error probability equivalent to the outage probability[3]. Therefore, the information-theoretic characterization consists of finding the outage probabilities associated with rate growths of r_a, r_b in the message sets $\mathcal{A}, \mathcal{B}$.

5.2. Achievable Region. Since different error probability behaviour (diversity order) for the two information streams is desired, we want to analyze the individual error probabilities induced by a coding and decoding strategy. Given a joint codebook $\mathbf{X}_{\mathbf{a},\mathbf{b}}$, indexed by the information from message sets $\mathcal{A}, \mathcal{B}$, and a joint decoding scheme $\mathcal{D}$, we want to characterize the individual error probabilities

[3]An alternate interpretation is that the diversity can be shown to be equivalent to the outage probability in the SNR exponent similar to that in [**24**].

on the message sets $\mathcal{A}$ and $\mathcal{B}$. More precisely, if the decoder outputs $\{\hat{\mathbf{a}}, \hat{\mathbf{b}}\}$, and the transmitted information is $\{\mathbf{a}, \mathbf{b}\}$, we want to characterize,

$$\mathbb{P}_e(\mathcal{A}) = \text{Prob}\left\{\{\hat{\mathbf{a}}, \hat{\mathbf{b}}\} : \hat{\mathbf{a}} \neq \mathbf{a}\right\}, \;\; \mathbb{P}_e(\mathcal{B}) = \text{Prob}\left\{\{\hat{\mathbf{a}}, \hat{\mathbf{b}}\} : \hat{\mathbf{b}} \neq \mathbf{b}\right\}. \tag{5.1}$$

In examining these error probabilities in terms of diversity order, we are interested in the high SNR regime. Instead of directly examining the error probabilities, we examine the outage behaviour of the codes with the understanding that this will closely reflect the error probability behaviour [**24**]. For different diversity orders desired of the message sets $\mathcal{A}$ and $\mathcal{B}$, the general characterization of the individual outage behaviour is a difficult problem. Given this, we take a simpler approach by examining a particular codebook design and decoding strategy.

We examine the additive codeword model introduced in Section 4.1, and define the power distribution between the two message sets. We generate a random codebook $\mathbf{X}_{\mathbf{a},\mathbf{b}} = \mathbf{X}_{\mathbf{a}} + \mathbf{X}_{\mathbf{b}}$ where both $\mathbf{X}_{\mathbf{a}}$ and $\mathbf{X}_{\mathbf{b}}$ are generated from i.i.d. Gaussian codebooks with average powers P_a, P_b. Given the total power constraint P, the power allocation needs to satisfy $P_a + P_b \leq P$. We use a successive decoding strategy, where $\mathbf{X}_{\mathbf{a}}$ is decoded considering $\mathbf{X}_{\mathbf{b}}$ as part of the noise. Then the decoded $\hat{\mathbf{X}}_{\mathbf{a}}$ is subtracted from the received signal $\mathbf{Y}$, and $\mathbf{X}_{\mathbf{b}}$ is decoded. Note that such a decoding strategy from a diversity order perspective may not be optimal, but is perhaps the simplest case to analyze. Given this coding and decoding strategy, one can write the individual outage probabilities $\mathbb{P}_{out}^{(\mathcal{A})}(\text{SNR}), \mathbb{P}_{out}^{(\mathcal{B})}(\text{SNR})$ as follows.

$$\begin{aligned} \mathbb{P}_{out}^{(\mathcal{A})}(\text{SNR}) &= \text{Prob}\left\{\mathbf{H} : \log \frac{|\mathbf{I} + \text{SNR}\mathbf{H}\mathbf{H}^*|}{|\mathbf{I} + \text{SNR}_b\mathbf{H}\mathbf{H}^*|} < R_a\right\} \\ \tilde{\mathbb{P}}^{(\mathcal{B})}(\text{SNR}) &= \text{Prob}\left\{\mathbf{H} : \log|\mathbf{I} + \text{SNR}_b\mathbf{H}\mathbf{H}^*| < R_b\right\} \\ \mathbb{P}_{out}^{(\mathcal{B})}(\text{SNR}) &= \max\left\{\mathbb{P}_{out}^{(\mathcal{A})}(\text{SNR}), \tilde{\mathbb{P}}^{(\mathcal{B})}(\text{SNR})\right\}, \end{aligned} \tag{5.2}$$

where $\text{SNR} = \frac{P}{\sigma^2}$, $\text{SNR}_b = \frac{P_b}{\sigma^2}$, and $|\mathbf{A}|$ denotes the determinant of the matrix $\mathbf{A}$. Note that the average achievable rate is given by,

$$R(SNR) = (1 - \mathbb{P}_{out}^{(\mathcal{A})}(\text{SNR}))R_a + (1 - \mathbb{P}_{out}^{(\mathcal{B})}(\text{SNR}))R_b, \tag{5.3}$$

which could be another criterion to optimize. We design the power allocation to the message sets $\mathcal{A}, \mathcal{B}$ respectively be such that $\frac{P_a}{P_b} = (P/\sigma^2)^\beta = \text{SNR}^\beta,\ \beta \in [0,1]$. Then, at high SNR, we have $\text{SNR}_a \doteq \text{SNR},\ \text{SNR}_b \doteq \text{SNR}^{1-\beta}$.

Then, for a given β, in a manner similar to [**24**], we can characterize $\mathbb{P}_{out}^{(\mathcal{A})}(\text{SNR})$ and $\tilde{\mathbb{P}}^{(\mathcal{B})}(\text{SNR})$, as

$$\mathbb{P}_{out}^{(\mathcal{A})}(\text{SNR}) \doteq \text{SNR}^{-\tilde{d}_a},\;\; \tilde{\mathbb{P}}^{(\mathcal{B})}(\text{SNR}) \doteq \text{SNR}^{-\tilde{d}_b}, \tag{5.4}$$

where

$$\begin{aligned} \tilde{d}_a &= \inf_{\alpha \in \mathcal{F}_a} \sum_{i=1}^{\min(M_t,M_r)} \left[|M_t - M_r| + 2i - 1\right]\alpha_i \\ \tilde{d}_b &= \inf_{\alpha \in \mathcal{F}_b} \sum_{i=1}^{\min(M_t,M_r)} \left[|M_t - M_r| + 2i - 1\right]\alpha_i, \end{aligned} \tag{5.5}$$

and

$$\mathcal{F}_a = \{\alpha_1 \geq \alpha_2 \ldots \geq \alpha_{\min(M_t,M_r)} \geq 0 : \sum_{i=1}^{\min(M_t,M_r)} (1-\alpha_i)^+ - \sum_{i=1}^{\min(M_t,M_r)} (1-\beta-\alpha_i)^+ < r_a\} \tag{5.6}$$

$$\mathcal{F}_b = \{\alpha_1 \geq \alpha_2 \ldots \geq \alpha_{\min(M_t,M_r)} \geq 0 : \sum_{i=1}^{\min(M_t,M_r)} (1-\beta-\alpha_i)^+ < r_b\},$$

where $(\cdot)^+ = \max(\cdot, 0)$. Since successive decoding is employed, the diversity levels achieved for the given β are given by $D_a(\beta) = \tilde{d}_a$ and $D_b(\beta) = \min(\tilde{d}_a, \tilde{d}_b)$, and the decoding order would matter. Therefore, the achievable tuple using this particular coding and decoding strategy is parameterized by β. This characterization is quite straightforward and for specific cases (such as $M_t = M_r = 2$) one can obtain explicit solutions to (5.5). However, this may not be a tight characterization of the achievable region since successive decoding may not be the right approach for diversity considerations, and the chosen power allocation strategy might not be optimal.

In more detail, for $M_t = 2 = M_r$, one can demonstrate that we can achieve the following rate tuple.

$$\begin{aligned}
\text{If } r_a \in (0,1), D_a &< 1 + 3(1-r_a) \\
\text{If } r_a \in (1,2), D_a &< 2 - r_a \\
\text{If } r_a \in (0,1), \text{and } (r_a + r_b) \in (0,1) & \\
D_b &< 4(1-r_a) - 3r_b = D^*(r_a+r_b) - r_a \\
\text{If } r_a \in (0,1), \text{and } (r_a + r_b) \in (1,2) & \\
D_b &< 2 - r_a - (r_a + r_b) = D^*(r_a+r_b) - r_a \\
\text{If } r_a \in (1,2), D_b &= 0,
\end{aligned} \tag{5.7}$$

where $D^*(r)$ denotes the optimal diversity order for a multiplexing gain of r as given in [**24**]. Therefore, atleast for the case of $M_t = M_r = 2$, one obtains a tuple (r_a, D_a, r_b, D_b) such that (r_a, D_a) lies on the optimal rate-diversity trade-off curve specified in [**24**] and (r_a+r_b, D_b) loses diversity order of exactly r_a over the optimal. Note that since we cannot dominate the "single-layer" rate-diversity tradeoff, we have a natural outer-bound for the rate tuple.

$$\begin{aligned}
D_a &< D^*(r_a) \\
D_b &< D^*(r_b) \\
\min(D_a, D_b) &< D^*(r_a + r_b).
\end{aligned} \tag{5.8}$$

Therefore, the best we could have hoped for when $M_t = 2 = M_r$ is for $(r_a + r_b, D_b)$ to lie on the optimal rate-diversity curve as specified in [**24**]. However, we are below it by an amount of r_a. It is unclear if this is due to the particular coding/decoding scheme used or if the "single-layer" outer bound is not achievable.

There might be a technique based on recent results on the multiple-antenna broadcast channel (see for example [**23**] among others) allowing us to employ a precoding strategy that is different from just superposition coding. The difference

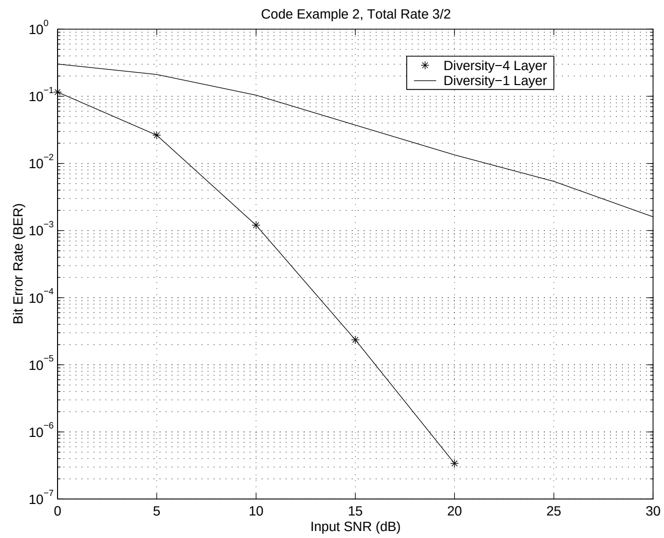

FIGURE 3. Error performance of Embedded code in Example 2

in our case from the equivalent broadcast channel considered there is that the channel is necessarily unknown at the transmitter for us, whereas it is assumed to be known in their case. However, we are currently investigating coding strategies that go beyond superposition coding to improve the inner bound.

6. Numerical Example

To illustrate the diversity-embedding property of the codes in Section IV, consider code example 2. Figure 3 depicts the bit error rate for the two information layers where the slopes of the 2 curves are markedly different indicating different diversity levels. We assume BPSK modulation (for simplicity), a Rayleigh flat-fading channel, and maximum likelihood decoding. At a BER of 10^{-2}, one layer can be decoded under much worse channel conditions than the other layer (about 15 dB less SNR). Another view point is the following : at given channel conditions (say an SNR of 15 dB), one layer can include information with more stringent reliability requirements (e.g. real-time data or video) than the other layer (e.g. voice or non-real-time data) and the 2 layers can still be decoded successfully meeting their individual QoS requirements.

7. Discussion

In this paper we introduced the idea of diversity embedded space-time codes in which a high-rate space-time code is designed to have a (lower-rate) high-diversity code embedded within it. The main idea was to opportunistically use a quasi-static channel realization without knowing it at the transmitter. The constructions given were better than simple time-sharing and we also explored an information-theoretic characterization of the achievable tuples. There are several open questions including characterization of the achievable region, extremal code constructions and practical performance under complexity constraints. We hope that this work is a first step towards understanding these issues.

ACKNOWLEDGEMENTS

We gratefully acknowledge several stimulating discussions with Vinay Vaishampayan and constructive comments from an anonymous reviewer.

References

[1] S.M. Alamouti. A simple transmit diversity technique for wireless communications. *IEEE Journal on Selected Areas in Communications*, 16(8):1451–1458, October 1998.

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

[3] Shlomo Shamai (Shitz). A broadcast strategy for the Gaussian slowly fading channel. In *IEEE International Symposium on Information Theory (ISIT)*, Ulm, Germany, p 150, June 1997.

[4] A. R. Calderbank and A. F. Naguib. Orthogonal designs and third generation wireless communication. In J. W. P. Hirschfeld, editor, *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] Thomas M. Cover. Broadcast channels. *IEEE Transactions on Information Theory*, 18:2–14, January 1972.

[7] Thomas M. Cover and Joy A. Thomas. *Elements of Information Theory.* John Wiley and Sons, Inc., New York, 1991.

[8] Imre Csiszar and Janos Korner. *Information Theory: Coding theorems for Discrete Memoryless Channels.* Academic, New York, 1982.

[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] M O. Damen, H. El-Gamal, and N. Beaulieu. On optimal linear space-time constellations. In *International Conference on Communications (ICC)*, May 2003.

[11] Suhas N. Diggavi. On achievable performance of spatial diversity fading channels. *IEEE Transactions on Information Theory*, 47(1):308–325, January 2001.

[12] 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.

[13] Jiann-Ching Guey, Michael P. Fitz, Mark R. Bell, and Wen-Yi Kuo. Signal design for transmitter diversity wireless communication systems over Rayleigh fading channels. *IEEE Transactions on Communications*, 47(4):527–537, April 1999.

[14] J. Hagenauer. Rate Compatable Punctured Convolutional Codes (RCPC codes) and their applications. *IEEE Transactions on Communications*, 36(4):389–400, April 1988.

[15] B. Hassibi and B. Hochwald. High-Rate Codes That Are Linear in Space and Time. *IEEE Transactions on Information Theory*, pages 1804–1824, July 2002.

[16] L H. Ozarow, S Shamai, and A. D. Wyner. Information theoretic considerations for cellular mobile radio. *IEEE Transactions on Vehicular Technology*, 43(2):359–378, May 1994.

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

[18] Shlomo Shamai (Shitz) and A. Steiner. Single user broadcasting in a MIMO channel. In *Information Theory Workshop (ITW2003)*, Paris, France, pp 38–41, April 2003.

[19] 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.

[20] 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.

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

[22] R. van Nobelen, N. Seshadri, J. Whitehead, and S. Timiri. An adaptive radio link protocol with enhanced data rates for GSM evolution. *IEEE Personal Communications*, 6(1), February 1999.

[23] W Yu. *Competition and Cooperation in Multiuser Communication Environments*. PhD thesis, Stanford University, Stanford, CA., 2002.

[24] Lizhong 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. Available from http://degas.eecs.berkeley.edu/ dtse/pub.html.

AT&T Shannon Labs, 180, Park Ave, Bldg 103, Florham Park, New Jersey, NJ 07932, USA.

Current address: School of Computer and Communication Sciences, Laboratory of Information and Communication Systems (LICOS), Swiss Federal Institute of Technology (EPFL), EPFL-IC-ISC-LICOS, Bldg, PSE A, Lausanne, Switzerland, CH 1015.

E-mail address: suhas.diggavi@epfl.ch

Department of Electrical Engineering University of Texas at Dallas, Richardson, TX 75083 USA.

E-mail address: aldhahir@utdallas.edu

Program in Applied and Computational Mathematics, Princeton University, Princeton, New Jersey, USA.

E-mail address: calderbank@math.princeton.edu