Capacity-Optimal Structured Linear Dispersion Codes for Correlated MIMO Channels

(Invited Paper)

Akbar M. Sayeed
ECE Department
University of Wisconsin-Madison
Madison, WI 53706
Email: akbar@engr.wisc.edu

Jayesh H. Kotecha
ECE Department
University of Wisconsin-Madison
Madison, WI 53706
Email: jkotecha@ece.wisc.edu

Zhihong Hong
Communications Research Centre Canada
Ottawa, Ontario K2H 8S2 Canada
Email: zhihong.hong@crc.ca

***Abstract*— We investigate the design of linear dispersion (LD) space-time codes for correlated MIMO channels. Our analysis is based on a unitarily equivalent eigen-domain representation of correlated MIMO fading channels. The eigen-domain channel matrix has statistically independent entries and the channel correlation structure is represented in the non-uniform powers of the entries. Capacity and pairwise error probability (PEP) analysis is greatly simplified in the eigen-domain. In particular, the capacity-achieving input covariance matrix is diagonal in the eigen-domain, and the PEP bounds reveal the interaction between code and the channel in spatio-temporal signal space dimensions. Furthermore, the achievable spatial multiplexing gain and diversity are constrained by the number of dominant channel entries in the eigen-domain which varies with SNR. Using insights from the capacity and PEP analysis, we propose a characterization of capacity-optimal LD codes via a family of structured code generator matrices. The family of generator matrices is parameterized by three unitary matrices that determine the space-time structure of the codes and a diagonal power-shaping matrix. The role of these matrices in controlling code performance is discussed and illustrative numerical results are presented.**

I. INTRODUCTION

Linear dispersion (LD) space-time codes were introduced in [1]. The key idea is to use space-time code matrices to linearly spread the information symbols in both space and time. LD codes are attractive from a signal space perspective since the information symbols can be drawn independently and redundancy (coding) is produced by appropriately choosing the space-time code matrices. The design approach in [1] was primarily based on maximizing the mutual information achieved by the code, and the design of LD codes based on pairwise error probability (PEP) optimization was investigated in [2]. LD codes can also be construed as lattice codes which have been recently shown in [3] to achieve the multiplexing-diversity tradeoff for i.i.d. MIMO channels [4]. The above schemes, as well as most existing space-time coding techniques, have been developed for an i.i.d. model for the MIMO channel that assumes a rich scattering environment. However, practical MIMO channels are seldom i.i.d. and exhibit spatially correlated fading. As a result, the performance of existing space-time codes designed for i.i.d. channels can degrade severely (see for example [5],[6], [7]) under correlated fading.

We investigate the design of LD codes for correlated MIMO Rayleigh fading channels in this paper. We assume that the channel is perfectly known at the receiver, but only channel covariance is known at the transmitter. When there is no channel information at the transmitter, a reasonable approach is to design robust codes for various channel conditions, [8], [7]. Our approach is based on a canonical eigen-domain channel representation [9], [10] that transforms an arbitrary correlated MIMO channel matrix into a unitarily equivalent eigen-domain matrix whose entries are statistically independent but not identically distributed; they have non-uniform powers associated with them. This channel representation is reviewed in the next section.

As discussed in Section III, the independence of channel entries in the eigen-domain greatly simplifies capacity and PEP analysis. In particular, the capacity-maximizing input covariance matrix is diagonal in the eigen-domain, and, unlike i.i.d. channels, the spatial multiplexing gain (rank of the optimal input) varies with SNR; it increases from 1 at low SNR (beamforming) to maximum at high SNR. Similarly, PEP analysis in the eigen-domain shows that the error rate of a code depends not only on the code but also on its interaction with the channel via the channel covariance matrix. In particular, the maximum achievable diversity gain is equal to the dominant non-vanishing channel entries in the eigen-domain, which also varies with SNR.

In Section IV, using insights from both capacity and PEP analysis, we develop a characterization of capacity-optimal LD codes based on a family of structured code generator matrices. The characterization is based on the idea of tensor product spaces and the family of generator matrices is parameterized by three unitary matrices, that determine the space-time code structure, and a diagonal power shaping matrix. The impact of these matrices on code performance as well as their design is discussed and illustrative numerical results are provided.

II. SYSTEM MODEL

Consider a narrowband frequency non-selective MIMO channel with M_t transmit and M_r receive antennas. With k indicating discrete time, if $\mathbf{s}(k)$ is the transmit vector of dimension M_t, then the M_r-dimensional received signal $\mathbf{x}(k)$

is given by

$$\mathbf{x}(k) = \sqrt{\frac{\mathcal{E}_s}{M_t}}\mathbf{H}(k)\mathbf{s}(k) + \mathbf{n}(k) \tag{1}$$

where $\mathbf{H}(k)$ is the $M_r \times M_t$ channel matrix coupling the transmitter and receiver antennas, and $\mathbf{n}(k)$ is the M_r-dimensional zero-mean, complex Gaussian white noise vector with unit variance; $E[\mathbf{n}(k)\mathbf{n}^H(k)] = \mathbf{I}_{M_r}$. $\mathcal{E}_s$ is the total transmit power (and denotes the normalized transmit SNR). The transmitted signal satisfies the power constraint $E[\|\mathbf{s}(k)\|^2] = M_t$. The channel covariance matrix is defined as $\mathbf{R} = E(\mathbf{h}\mathbf{h}^H)$, where $\mathbf{h} = \mathrm{vec}(\mathbf{H})$. The channel matrix is assumed known at the receiver but not at the transmitter. However, it is assumed that the transmitter knows the channel covariance matrix $\mathbf{R}$.

We consider a block fading MIMO channel; that is, $\mathbf{H}(k) = \mathbf{H}$ for $k = 1, \ldots, T$, and the channel is independent between different blocks of T symbols. We consider a single coherence interval of duration $T \geq M_t$ and suppress the index k in $\mathbf{H}(k)$. Thus, the system equation over one block is

$$\mathbf{X} = \sqrt{\frac{\mathcal{E}_s}{M_t}}\mathbf{H}\mathbf{S} + \mathbf{N}\ , \tag{2}$$

where $\mathbf{X} = [\mathbf{x}(1), \ldots, \mathbf{x}(T)]$ is the $M_r \times T$ received signal matrix, $\mathbf{S} = [\mathbf{s}(1), \ldots, \mathbf{s}(T)]$ is the $M_t \times T$ transmitter signal matrix, and $\mathbf{N} = [\mathbf{n}(1), \ldots, \mathbf{n}(T)]$. The transmitted signal matrix satisfies the power constraint $E[\mathrm{trace}[\mathbf{S}\mathbf{S}^H]] = M_t T$. Maximum likelihood decoding is assumed at the receiver.

A. Channel Model

Let $\mathbf{\Sigma}_t = E[\mathbf{H}^H\mathbf{H}]$ and $\mathbf{\Sigma}_r = E[\mathbf{H}\mathbf{H}^H]$ denote the transmit and receive channel covariance matrices, respectively. Furthermore, let $\mathbf{\Sigma}_t = \mathbf{U}_t\mathbf{\Lambda}_t\mathbf{U}_t^H$ and $\mathbf{\Sigma}_r = \mathbf{U}_r\mathbf{\Lambda}_r\mathbf{U}_r^H$ denote their eigendecompositions, where $\mathbf{U}_t$ and $\mathbf{U}_r$ are unitary matrices of eigenvectors (eigenmatrices), and $\mathbf{\Lambda}_t$ and $\mathbf{\Lambda}_r$ denote the corresponding diagonal matrices of eigenvalues. It is shown in [10] (see also [11]) that under mild assumptions the channel matrix admits the following canonical Karhunen-Loeve (K-L) expansion[1]

$$\mathbf{H} = \mathbf{U}_r\mathbf{H}_c\mathbf{U}_t^H \quad , \quad \mathbf{H}_c = \mathbf{U}_r^H\mathbf{H}\mathbf{U}_t \tag{3}$$

where the canonical channel matrix $\mathbf{H}_c$ has uncorrelated (independent for Rayleigh fading) entries but they are not identically distributed in general (the variances can be different). Thus, the K-L representation (3) transforms an arbitrary correlated channel matrix $\mathbf{H}$ into a unitarily equivalent channel matrix $\mathbf{H}_c$ with uncorrelated entries. The channel covariance matrix admits the following eigen-decomposition

$$\mathbf{R} = E[\mathbf{h}\mathbf{h}^H] = (\mathbf{U}_t^* \otimes \mathbf{U}_r)\mathbf{R}_c(\mathbf{U}_t^* \otimes \mathbf{U}_r)^H \tag{4}$$

where $\otimes$ denotes the kronecker/tensor product and $\mathbf{R}_c = E[\mathbf{h}_c\mathbf{h}_c^H]$ is the *diagonal* covariance matrix of the eigen-domain matrix $\mathbf{H}_c$, with $\mathbf{h}_c = \mathrm{vec}(\mathbf{H}_c)$.

Two important channel models can be expressed in the form (3). In [9], a *virtual* channel representation is proposed for uniform linear arrays (ULA) of antennas. It is shown that the channel in this case can be written in the form (3), where $\mathbf{U}_r$ and $\mathbf{U}_t$ are discrete Fourier transform (DFT) matrices. In the case of the popular kronecker channel model, $\mathbf{H} = \mathbf{\Sigma}_r^{1/2}\mathbf{H}_w\mathbf{\Sigma}_t^{1/2} = \mathbf{U}_r\mathbf{\Lambda}_r^{1/2}\mathbf{H}_w\mathbf{\Lambda}_t^{1/2}\mathbf{U}_t^H$ where $\mathbf{H}_w$ is an i.i.d. matrix, and the second equality corresponds to the canonical form (3) with $\mathbf{H}_c = \mathbf{\Lambda}_r^{1/2}\mathbf{H}_w\mathbf{\Lambda}_t^{1/2}$. For the kronecker channel, $\mathbf{R}_c = \mathbf{\Lambda}_t \otimes \mathbf{\Lambda}_r$ which is a special case of (4) and clearly shows the separability in the transmit and receive correlation structure implicit in the kronecker model. In general, correlated MIMO channels do not admit this separable structure and the eigen-decomposition (4) represents the most general form.

The uncorrelated nature of $\mathbf{H}_c$ suggests signaling and and reception in the eigen-domain:

$$\mathbf{X}_c = \mathbf{H}_c\mathbf{S}_c + \mathbf{N}_c \tag{5}$$

where $\mathbf{S}_c = \mathbf{U}_t^H\mathbf{S}$ and $\mathbf{X}_c = \mathbf{U}_r^H\mathbf{X}$. Note that (5) is a unitarily equivalent representation of the channel and thus all aspects of system design, such as capacity and PEP analysis, can be directly done in the eigen-domain. The transmitted signal $\mathbf{S}_c$ is modulated onto the transmit eigenfunctions (as $\mathbf{S} = \mathbf{U}_t\mathbf{S}_c$) before launching into the channel, and the received signal is first projected onto the receive eigenfunctions before decoding. For the remainder of the paper, we work in the eigen-domain (5) and suppress the subscript 'c' unless necessary.

III. Performance Metrics

A. Capacity

It has been shown in [12] that the ergodic capacity of a Rayleigh fading MIMO channel is achieved by a zero mean, complex, Gaussian input. Let $\mathbf{Q} = E[\mathbf{s}\mathbf{s}^H]$ be the covariance matrix of the input in the eigendomain. It is shown in [10] that the optimal input covariance matrix is diagonal in the eigendomain, that is $\mathbf{Q}_{opt} = \mathbf{\Lambda}_{opt}$, and solves

$$C = \max_{\mathbf{\Lambda}:\mathrm{tr}(\mathbf{\Lambda})\leq M_t} E_{\mathbf{H}}\left[\log\left|\mathbf{I} + \frac{\mathcal{E}_s}{M_t}\mathbf{H}\mathbf{\Lambda}\mathbf{H}^H\right|\right]. \tag{6}$$

The diagonal entries of $\mathbf{\Lambda}_{opt}$ reflect the power allocated to different transmit eigendimensions to achieve capacity and they depend on the SNR. Let $d(\mathrm{SNR}) = \mathrm{rank}(\mathbf{\Lambda}_{\mathrm{opt}}(\mathrm{SNR}))$ denote the number of transmit eigendimensions excited at a given SNR. Note that $1 \leq d(\mathrm{SNR}) \leq \mathrm{M_t}$ reflects the spatial multiplexing gain provided by the channel. It is well-known that, for an i.i.d. channel, uniform power allocation is optimal for all SNRs. In contrast, for correlated channels, non-uniform allocation is needed in general. In particular, beamforming is optimal at sufficiently low SNR ($d = 1$). As the SNR increases, $d(\mathrm{SNR})$ increases, and at high SNR $d = M_t$ and $\mathbf{\Lambda}_{opt} \rightarrow \mathbf{I}$ as in an i.i.d. channels. Thus, in contrast to i.i.d. channels, the spatial multiplexing gain ($d(\mathrm{SNR})$) is an increasing function of SNR. The structured LD codes proposed in Section IV exploit this property of optimal input signal.

[1] Best evident in the vectorized form: $\mathbf{h} = [\mathbf{U}_t^* \otimes \mathbf{U}_r]\mathbf{h}_c$ [9].

B. Pairwise Error Probability

Let $\mathbf{E} = \mathbf{S} - \hat{\mathbf{S}}$ denote the error codeword matrix in the eigendomain, where $\mathbf{S}$ and $\hat{\mathbf{S}}$ are the transmitted and decoded codeword matrices. Denote the codeword error covariance matrix as $\mathbf{R}_e = \mathbf{E}\mathbf{E}^H$. Let $P(\mathbf{S} \to \hat{\mathbf{S}}|\mathbf{H})$ denote the probability of decoding $\hat{\mathbf{S}}$ when $\mathbf{S}$ is transmitted, conditioned on the channel realization $\mathbf{H}$. Then, the pairwise error probability (PEP) can be upper bounded as

$$\begin{aligned} P(\mathbf{S} \to \hat{\mathbf{S}}) &= E_{\mathbf{H}}\left(P(\mathbf{S} \to \hat{\mathbf{S}}|\mathbf{H})\right) \\ &\leq \left|\mathbf{I}_{M_t M_r} + \frac{\mathcal{E}_s}{4M_t}\mathbf{R}(\mathbf{R}_e \otimes \mathbf{I}_{M_r})\right|^{-1} \quad (7) \\ &= \left|\mathbf{I}_{M_t M_r} + \frac{\mathcal{E}_s}{4M_t}\tilde{\mathbf{R}}(\mathbf{I}_{M_r} \otimes \mathbf{R}_e)\right|^{-1} \quad (8) \end{aligned}$$

where $\tilde{\mathbf{R}}$ corresponds to the covariance of the row stacked vector of $\mathbf{H}$. Note that both $\mathbf{R}$ and $\tilde{\mathbf{R}}$ are diagonal in the eigendomain. The PEP depends on the interaction between the codewords and the channel that is captured by the $M_t M_r \times M_t M_r$ matrix $\mathbf{\Delta} = \tilde{\mathbf{R}}(\mathbf{I}_{M_r} \otimes \mathbf{R}_e)$:

$$\mathbf{\Delta} = \tilde{\mathbf{R}}(\mathbf{I} \otimes \mathbf{R}_e) = \begin{bmatrix} \tilde{\mathbf{R}}(1)\mathbf{R}_e & & \\ & \ddots & \\ & & \tilde{\mathbf{R}}(M_r)\mathbf{R}_e \end{bmatrix} \quad (9)$$

where $\tilde{\mathbf{R}}(i)$ is the diagonal covariance matrix of the i-th row of $\mathbf{H}$, representing the MISO channel from all transmit eigen-dimensions to the i-th receive eigen-dimension. Thus, the PEP bound decomposes as

$$\begin{aligned} \text{PEP} &\leq \prod_{i=1}^{M_r}\left|\mathbf{I}_{M_r} + \frac{\mathcal{E}_s}{4M_t}\tilde{\mathbf{R}}(i)\mathbf{R}_e\right|^{-1} \\ &= \prod_{i=1}^{M_r}\prod_{j=1}^{M_t}\left(1 + \frac{\mathcal{E}_s}{4M_t}\lambda_j(\tilde{\mathbf{R}}(i)\mathbf{R}_e)\right)^{-1} \quad (10) \end{aligned}$$

where $\lambda_j(\tilde{\mathbf{R}}(i)\mathbf{R}_e)$ denotes the j-the eigenvalue of $\tilde{\mathbf{R}}(i)\mathbf{R}_e$. Let $d_i = \text{rank}(\tilde{\mathbf{R}}(i)),\ i = 1, \cdots, M_r$, denote the number of non-zero diagonal elements (eigenvalues) of $\tilde{\mathbf{R}}(i)$. Note that $d_i \leq M_t$ and $\text{rank}(\mathbf{R}_e) \leq M_t$. The achievable diversity gain, div, the number of non-zero eigenvalues of $\mathbf{\Delta}$, is bounded as

$$0 \leq \text{div} = \sum_{i=1}^{M_r}\text{rank}(\tilde{\mathbf{R}}(i)\mathbf{R}_e) \leq \sum_{i=1}^{M_r} d_i = D \leq M_r M_t \quad (11)$$

where the upperbound D corresponds to full-rank $\mathbf{R}_e$. Note that $D = \text{rank}(\tilde{\mathbf{R}})$ is the number of non-zero diagonal elements in $\tilde{\mathbf{R}}$ (the number of entries in $\mathbf{H}_c$ with non-vanishing variance) and represents the total degrees of freedom in the correlated channel. The coding gain is given by the product of the non-zero eigenvalues of $\mathbf{\Delta}$. For BLAST-type signaling, $\text{rank}(\mathbf{R}_e) \leq 1$ and it follows from (11) that $\text{div} \leq M_r$, whereas for space-time coding with full-rank $\mathbf{R}_e$, $\text{div} = D$.

Let $d_r = \text{rank}(\mathbf{\Lambda}_r) \leq M_r$ denote the number of non-zero (diagonal) entries in $\mathbf{\Lambda}_r$ (receive channel covariance matrix), and let $d_t = \text{rank}(\mathbf{\Lambda}_t) \leq M_t$ denote the number of non-zero entries in $\mathbf{\Lambda}_t$. It follows that $D \leq d_t d_r$ and the maximum multiplexing gain is $\min(d_r, d_t) \leq \min(M_r, M_t)$. A maximum of d_t transmit eigen dimensions need to be excited at high SNR (corresponding to the d_t columns of $\mathbf{H}_c$ with at least one entry with non-vanishing power). This results in a smaller multiplexing gain (compared to i.i.d. channels) if $d_t < M_t$ but an increased coding gain since the total transmit power is concentrated in fewer transmit dimensions. We assume that $d_r = M_r$ and $d_t = M_t$ but $D < M_r M_t$ (sparse $\mathbf{R}$; see (11)).

IV. Structured Linear Dispersion Codes

In this section, we introduce a framework for designing structured LD codes for correlated MIMO channels based on the concept of tensor products spaces. Our design is motivated by both capacity and PEP analyses.

We consider LD codes in the eigendomain (and suppress the subscript c) of the form [1]

$$\mathbf{S} = \sum_{n=1}^{N}\mathbf{A}_n a_n \quad (12)$$

where $\{a_n\}_{n=1}^N$ are i.i.d. information symbols from a complex constellation of size Q, and $\{\mathbf{A}_n\}_{n=1}^N$ is a set of $M_t \times T$ complex space-time codeword matrices that define the code.[2] It is well-known that the capacity of a MIMO channel at high SNR scales as $C \sim \min(M_r, M_t)\log(\text{SNR})$ bits/s. Thus, over a block length of T seconds $CT \sim T\min(M_r, M_t)\log(\text{SNR})$ bits should be communicated when operating near capacity. The number of codeword matrices $N \leq \min(M_r, M_t)T = N_o$ corresponds to the number of space-time signal space dimensions utilized for communication and the constellation size, Q, for $\{a_n\}$ is chosen corresponding to the SNR. The rate of the LD code is

$$R = \frac{N}{T}\log_2(Q) \quad \text{bits/s.} \quad (13)$$

We assume that $M_r \geq M_t$ and $T \geq M_t$. In this case, chosing $N = M_t T = N_o$ results in a *full-rate* code [1] in the sense that the maximum available spatio-temporal signal space dimensions are used for communication.

The structure of LD codes can be studied by writing (5)in a column-stacked form

$$\mathbf{x} = \sqrt{\frac{\mathcal{E}_s}{M_t}}(\mathbf{I}_T \otimes \mathbf{H})\mathbf{s} + \mathbf{n} = \sqrt{\frac{\mathcal{E}_s}{M_t}}\mathcal{H}\mathbf{G}\mathbf{a} + \mathbf{n} \quad (14)$$

where $\mathcal{H} = \mathbf{I}_T \otimes \mathbf{H}$, $\mathbf{x} = \text{vec}(\mathbf{X})$ is $M_r T$ dimensional, $\mathbf{s} = \text{vec}(\mathbf{S}) = \mathbf{G}\mathbf{a}$ is $M_t T$ dimensional, $\mathbf{a} = [a_1, a_2, \cdots, a_N]^T$, and $\mathbf{G} = [\mathbf{g}_1, \cdots, \mathbf{g}_N] = [\text{vec}(\mathbf{A}_1), \text{vec}(\mathbf{A}_2), \cdots, \text{vec}(\mathbf{A}_N)]$ is an $N_o \times N$ *code generator matrix* that is an equivalent representation of the LD code. The input signal $\mathbf{s}$ in (14) satisfies the power constraint $E[\|\mathbf{s}\|^2] = M_t T$. We assume a unit power constraint on the i.i.d. information symbols; $E[|a_n|^2] = 1$ and $E[\mathbf{a}\mathbf{a}^H] = \mathbf{I}_N$. Then, the transmit power

[2] Note that our development can be extended to LD codes where the real and imaginary symbol components are modulated by separate matrices.

constraint can be expressed in terms of the code matrices as: $\text{trace}(\mathbf{G}\mathbf{G}^H) = \sum_n \text{trace}(\mathbf{A}_n\mathbf{A}_n^H) = M_t T$.

Let $\mathbf{Q} = E[\mathbf{s}\mathbf{s}^H]$ denote the covariance matrix of $\mathbf{s}$ in (14). Using (14) and (6), it can be readily shown (see also (16)) that the capacity-optimal $\mathbf{Q}$ for the vectorized channel is block diagonal and is given by $\mathbf{Q}_{opt} = \mathbf{I}_T \otimes \mathbf{\Lambda}_{opt}$. Now, for an LD code, the statistics of $\mathbf{s}$ are induced by $\mathbf{G}$ since $\mathbf{s} = \mathbf{G}\mathbf{a}$; that is, $\mathbf{Q} = \mathbf{G}\mathbf{G}^H$ since $E[\mathbf{a}\mathbf{a}^H] = \mathbf{I}$. We immediately have the following result.

Theorem 1: An LD code is capacity-optimal if its generator matrix satisfies

$$\mathbf{G}\mathbf{G}^H = \mathbf{I}_T \otimes \mathbf{\Lambda}_{opt}. \tag{15}$$

Another way to see this result is that, under the assumption of i.i.d, unit variance Gaussian information symbols $\{a_n\}$, the mutual information achieved by the LD code is

$$C(\mathbf{G}) = \frac{1}{T} E_{\mathbf{H}} \log \left| \mathbf{I}_{M_r T} + \frac{\mathcal{E}_s}{M_t} \mathcal{H}\mathbf{G}\mathbf{G}^H\mathcal{H}^H \right|. \tag{16}$$

which precisely equals (6) when $\mathbf{G}$ satisfies (15).[3]

A. Code Design Using Tensor Product Spaces

How do we design $\mathbf{G}$ to satisfy (15)? How do we control its performance in terms of probability of error? To address these questions, we study the structure of $\mathbf{G}$:

$$\mathbf{G} = \begin{bmatrix} \mathbf{G}_1 \\ \vdots \\ \mathbf{G}_t \\ \vdots \\ \mathbf{G}_T \end{bmatrix} = \begin{bmatrix} \mathbf{g}_{11} & \cdots & \mathbf{g}_{1n} & \cdots & \mathbf{g}_{1N} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ \mathbf{g}_{t1} & \cdots & \mathbf{g}_{tn} & \cdots & \mathbf{g}_{tN} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ \mathbf{g}_{T1} & \cdots & \mathbf{g}_{Tn} & \cdots & \mathbf{g}_{TN} \end{bmatrix} \tag{17}$$

where each $\mathbf{G}_t$ is $M_t \times N$ and each $\mathbf{g}_{tn}$ is M_t dimensional. The n-th column of $\mathbf{G}$ corresponds to $\mathbf{A}_n$ in (12)

$$\mathbf{g}_n = \text{vec}(\mathbf{A}_n) \quad , \quad \mathbf{A}_n = [\mathbf{g}_{1n}, \cdots, \mathbf{g}_{tn}, \cdots, \mathbf{g}_{Tn}] \tag{18}$$

and $\mathbf{S}$ in (12) can be alternatively expressed as (also used in [2])

$$\mathbf{S} = [\mathbf{G}_1\mathbf{a}, \cdots, \mathbf{G}_2\mathbf{a}, \cdots, \mathbf{G}_T\mathbf{a}] \tag{19}$$

Thus, LD codes can be thought of as a collection of N matrices in $\mathcal{C}^{M_t \times T}$ or as a collection of N vectors in $\mathcal{C}^{M_t T}$. Both spaces are of dimension $N_o = M_t T$ and we exploit this connection to design LD codes using tensor product spaces.

We illustrate the basic idea for full-rate codes so that $N = N_o = M_t T$ and $\mathbf{G}$ is $N_o \times N_o$. Let $\mathbf{u} \in \mathcal{C}^{M_t}$ and $\mathbf{v} \in \mathcal{C}^T$. Then $\mathbf{g} = \mathbf{v} \otimes \mathbf{u} \in \mathcal{C}^{M_t T}$ which could serve as a column of $\mathbf{G}$. The corresponding $M_t \times T$ matrix $\mathbf{A} = \mathbf{u}\mathbf{v}^T$. Note that $\mathbf{g} = \text{vec}(\mathbf{A})$ using the identity $\text{vec}(\mathbf{ADB}) = (\mathbf{B}^T \otimes \mathbf{A})\text{vec}(\mathbf{D})$. Now let $\mathbf{U} = [\mathbf{u}_1, \cdots, \mathbf{u}_{M_t}]$ be an $M_t \times M_t$ unitary matrix whose columns serve as an orthonormal basis (ONB) for $\mathcal{C}^{M_t}$. Similarly, let $\mathbf{V} = [\mathbf{v}_1, \cdots, \mathbf{v}_T]$ be a $T \times T$ unitary matrix. Then, the matrix

$$\mathbf{G} = \mathbf{V} \otimes \mathbf{U} = [\mathbf{v}_1 \otimes \mathbf{U}, \cdots, \mathbf{v}_t \otimes \mathbf{U}, \cdots, \mathbf{v}_T \otimes \mathbf{U}] \tag{20}$$

is an $N_o \times N_o$ unitary matrix whose columns/rows serve as an ONB for $\mathcal{C}^{M_t T}$. The corresponding code matrices $\{\mathbf{A}_n, n = 1, \cdots, M_t T\}$ can be extracted from the columns of $\mathbf{G}$ via (18):

$$\mathbf{A}_{t,m} = \mathbf{u}_m \mathbf{v}_t^T, \; t = 1, \cdots, T, \; m = 1, \cdots, M_t. \tag{21}$$

Thus, the columns of $\mathbf{V}$ provide temporal spreading, whereas the columns of $\mathbf{U}$ provide spatial spreading. The next result provides a characterization of structured capacity-optimal LD codes.

Theorem 2: Let $\mathbf{U}$ ($M_t \times M_t$), $\mathbf{V}$ ($T \times T$), and $\mathbf{\Phi}$ ($N_o \times N_o$) be arbitrary unitary matrices and let $\mathbf{\Lambda}_{opt}$ be the optimal input covariance matrix defined in (6). Any generator matrix of the form

$$\mathbf{G} = [\mathbf{V} \otimes \mathbf{\Lambda}_{opt}^{1/2}\mathbf{U}]\mathbf{\Phi} \tag{22}$$

defines a capacity-optimal, full-rate LD code.

The proof directly follows from the fact that $\mathbf{G}\mathbf{G}^H = \mathbf{V}\mathbf{V}^H \otimes \mathbf{\Lambda}_{opt}^{1/2}\mathbf{U}\mathbf{U}^H\mathbf{\Lambda}_{opt}^{1/2} = \mathbf{I}_T \otimes \mathbf{\Lambda}_{opt}$. The spatio-temporal signal space structure of the LD code is determined by $\mathbf{U}$, $\mathbf{V}$ and $\mathbf{\Phi}$: $\mathbf{U}$ provides spatial spreading, $\mathbf{V}$ temporal spreading and $\mathbf{\Phi}$ joint spatio-temporal spreading of information symbols. In particular, $\mathbf{\Phi}$ is critical to maximizing the diversity gain. The matrix $\mathbf{\Lambda}_{opt}$ is responsible for spatial power shaping; it does not affect the diversity gain but improves the coding gain. The codeword error matrix takes the following form for LD codes:

$$\mathbf{R}_e = \mathbf{E}\mathbf{E}^H = \sum_{t=1}^{T} \mathbf{G}_t \epsilon \epsilon^H \mathbf{G}_t^H = \sum_{n=1}^{N_o} \sum_{n'=1}^{N_o} \mathbf{A}_n \epsilon_n \epsilon_{n'}^* \mathbf{A}_{n'}^H \tag{23}$$

where $\{\mathbf{G}_t\}$ are defined in (17) and $\epsilon = \mathbf{a} - \hat{\mathbf{a}}$, the information symbol error vector. We elaborate by considering three progressive cases.

Case I: No precoding ($\mathbf{\Phi} = \mathbf{I}$) and no power shaping ($\mathbf{\Lambda}_{opt} = \mathbf{I}$). This choice of power shaping applies to i.i.d. channels or correlated channels at high SNR. In this case, the code matrices $\mathbf{A}_n$ are all rank-1 (see (21)). This is analogous to BLAST-type signaling and thus the diversity gain is limited to $\text{div} \leq M_r$. In particular, an error codeword in which only one ϵ_n is non-zero results in an rank-1 $\mathbf{R}_e = \mathbf{A}_n\mathbf{A}_n^H$.

Case II: Precoding ($\mathbf{\Phi} \neq \mathbf{I}$) and no power shaping ($\mathbf{\Lambda}_{opt} = \mathbf{I}$). An appropriate choice of the space-time precoding matrix $\mathbf{\Phi}$ can increase the rank of the spreading matrices $\{\mathbf{A}_n\}$ compared to Case I. This results in higher ranks for $\mathbf{R}_e$ compared to Case I, thereby increasing diversity gain to up to a maximum of D (see Section III-B). This is also noted in [1] where an explicit full-rate LD code, with full-rank orthogonal matrices ($\mathbf{A}_n\mathbf{A}_n^H = \mathbf{I}/M_t$), is derived (for the case $T = M_t$) via a unitary transformation ($\mathbf{\Phi}$) of rank-1 V-BLAST code matrices ($\mathbf{V} = \mathbf{U} = \mathbf{I}$) (see Eq. (31) in [1]).[4]. Thus, joint spatio-temporal spreading produced by $\mathbf{\Phi}$ in (22) plays a key role in increasing the diversity gain and hence improving the probability of error performance.

Case III: Precoding ($\mathbf{\Phi} \neq \mathbf{I}$) and optimal power shaping ($\mathbf{\Lambda}_{opt}$ as defined in (6)). This case provides capacity-optimal

[3]In ([1]), $C(\mathbf{G})$ is used as a metric for designing LD codes for i.i.d. channels.

[4]We note that the proposed structured LD codes in (22) subsume the codes presented in [1]

spatial power allocation to further improve the coding gain compared to Case II. Recall that $\mathbf{\Lambda}_{opt}$ is a function of SNR and approaches $\mathbf{I}$ at high SNR. Thus, we expect Case II and Case III to perform similarly at high SNRs.

Design of U, V, Φ. These three unitary matrices provide a rich space for optimizing the probability of error performance of a capacity-optimal LD code. Any design method may be used to design them (see [1] and [2], e.g.). One simple recipe for generating a $K \times K$ unitary matrix $\mathbf{U}$ is

$$\mathbf{U} = \mathbf{D}_K \text{diag}(1, e^{j\theta_1}, \cdots, e^{j\theta_{K-1}})\mathbf{D}_K \tag{24}$$

where $\mathbf{D}_K$ is any fixed $K \times K$ unitary matrix (e.g., a discrete Fourier transform matrix), and the phases $\{\theta_1, \cdots, \theta_{K-1}\}$ are chosen from $[0, 2\pi]$ to generate different $\mathbf{U}$'s.

B. Numerical Results

We present numerical results for full-rate LD codes for $M_t = M_r = T = 2$ ($N = N_o = 4$) to illustrate the three cases above: **Case I:** (no precoding), **Case II:** (with precoding), and **Case III:** (with precoding and power-shaping). The information symbols are independently drawn BPSK. A 2x2 correlated channel is considered whose entries in the eigendomain have variances $E[|H_c(1,1)|^2] = 1$, $E[|H_c(2,1)|^2] = 1$, $E[|H_c(1,2)|^2] = 0$, $E[|H_c(2,2)|^2] = 0.5$. Thus, $d_1 = 1$, $d_2 = 2$ and $D = 3$ (see (11)). For each case, three codes are compared. **Code 1**: $\mathbf{G}_{HH}$, based on the codes proposed in [1]. For Case I, $\mathbf{G}_{HH,I} = \mathbf{\Phi}_1$ corresponding to the unitary code matrices in Eq. (31) of [1]. For Case II, $\mathbf{G}_{HH,II} = \mathbf{G}_{HH,I}\mathbf{\Phi}_2$ where $\mathbf{\Phi}_2$ is a diagonal unitary matrix given in Eq. (34) in [1]. Finally, $\mathbf{G}_{HH,III} = (\mathbf{I} \times \mathbf{\Lambda}_{opt})\mathbf{G}_{HH,II}$ where $\mathbf{\Lambda}_{opt}$ is defined in (6) for the above channel. **Code 2**: $\mathbf{G}_{rand}$, based on (24) for generating random unitary matrices using DFT matrices. $\mathbf{G}_{rand,I} = \mathbf{V} \times \mathbf{U}$, $\mathbf{G}_{rand,II} = \mathbf{G}_{rand,I}\mathbf{\Phi}$, and $\mathbf{G}_{rand,III} = (\mathbf{I} \otimes \mathbf{\Lambda}_{opt})\mathbf{G}_{rand,II}$. **Code 3**: $\mathbf{G}_{opt}$, an optimized LD code used only in Case III. $\mathbf{G}_{opt} = (\mathbf{I} \otimes \mathbf{\Lambda}_{opt})\mathbf{G}_{HH,I}\mathbf{\Phi}_3$ where $\mathbf{\Phi}_3$ is a diagonal unitary matrix. The phases of $\mathbf{\Phi}_3$ were optimized to maximize $\delta = \min |\mathbf{R}_e|$ where the minimum is over all pairwise error vectors.

The symbol error probability for the three codes is shown in Fig. 1. In Case I, surprisingly, $\mathbf{G}_{rand,I}$ achieves a higher diversity gain of 2 even though it has rank-1 code matrices (BLAST-type) compared to the full-rank matrices of $\mathbf{G}_{HH,I}$. This is because the worst-case symbol error vector for $\mathbf{G}_{HH,I}$ has both entries non-zero, which results in rank reduction of $\mathbf{R}_e$. In this case, $\delta = \min |\mathbf{R}_e| = 0$ for both codes. Thus, higher rank spreading matrices are not better necessarily! In Case II, $\mathbf{G}_{HH,II}$ achieves the maximum diversity gain of 3 and $\delta = 0.9$ due to precoding, whereas $\mathbf{G}_{rand,II}$ still only achieves a diversity gain of 2 because δ is still 0. This is due to the DFT matrices used in $\mathbf{G}_{rand}$ (completely randomly chosen unitary matrices can yield minimum $\delta > 0$). Finally, in Case III, the performance of both $\mathbf{G}_{HH,II}$ and $\mathbf{G}_{rand,II}$ improves over Case II due to optimal power shaping (coding gain). However, the markedly improved performance is due to the optimized Code 3, $\mathbf{G}_{opt}$ which achieves $\delta = 4$, resulting in a coding gain of about 2dB over $\mathbf{G}_{HH,III}$ at $P_e = 10^{-4}$.

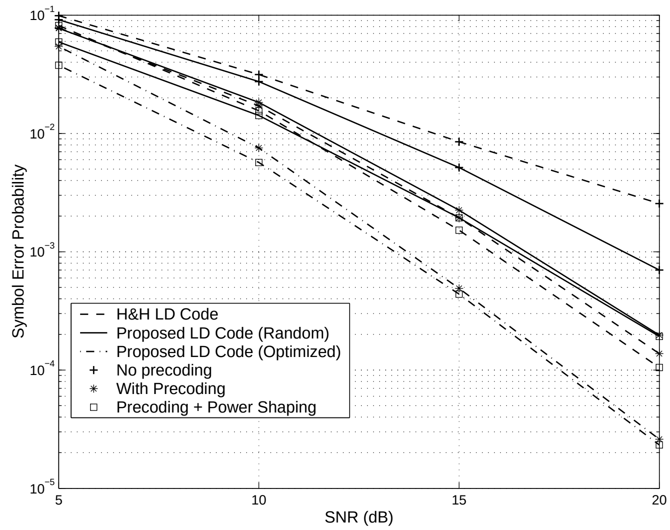

Fig. 1. *Performance comparison of the three codes.*

V. Conclusion

We have proposed a structured characterization of capacity-optimal LD codes for correlated MIMO channels. The design of the unitary matrices $\mathbf{U}$, $\mathbf{V}$ and $\mathbf{\Phi}$ to optimize the probability of error performance is a key challenge; the choice of $\mathbf{\Phi}$ is critical in this regard. We are currently investigating structured design of these matrices by leveraging the insights provided by the PEP analysis. This line of attack may prove fruitful in characterizing the multiplexing-diversity tradeoff [4] in correlated MIMO channels.

References

[1] B. Hassibi and B. Hochwald, "Optimal training in space-time systems," *Conference Record of the Thirty-Fourth Asilomar Conference on Signals, Systems and Computers*, vol. 1, pp. 743 –747, 2000.

[2] R. W. Heath and A. J. Paulraj, "Linear dispersion codes based on frame theory," *IEEE Trans. Signal Processing*, pp. 2429–2441, Oct. 2002.

[3] H. El Gamal, G. Gaire, and M. Damen, "Lattice coding and decoding achieve the optimal diversity-multiplexing tradeoff of MIMO channels," *IEEE Trans. Inform. Th.*, pp. 968–985, June 2004.

[4] L. Zheng and D. Tse, "Diversity and multiplexing: A fundamental tradeoff in multiple-antenna channels," *IEEE Trans. Inform. Th.*, pp. 1073–1096, May 2003.

[5] M. O. Damen, A. Abdi, and M. Kaveh, "On the effect of fading on several space-time coding and detection schemes," *IEEE Vehicular Technology Conference, Oct. 2001*, vol. 1, pp. 13–16.

[6] M. Uysal and C. N. Georghiades, "Effect of spatial fading correlation on performance of space-time codes," *IEEE Electronics Letters*, vol. 37, no. 3, pp. 181–183, Feb 2001.

[7] S. Siwamogsatham and M. P. Fitz, "Robust space-time coding for correlated Rayleigh fading channels," *IEEE Trans. of Signal Processing*, pp. 2408–2416, Oct. 2002.

[8] H. El Gamal, "On the robustness of space-time coding," *IEEE Trans. on Signal Processing*, pp. 2417–2428, Oct. 2002.

[9] A. M. Sayeed, "Deconstructing multi-antenna fading channels," *IEEE Trans. Signal Processing*, pp. 2563–2579, Oct. 2002.

[10] J. H. Kotecha and A. M. Sayeed, "Canonical statistical models for correlated MIMO fading channels and capacity analysis," *submitted to the IEEE Trans. Wireless Commun.*, 2004.

[11] J. H. Kotecha and A. M. Sayeed, "Transmit signal design for optimal estimation of correlated MIMO channels," *IEEE Trans. Signal Processing*, pp. 546–557, Feb. 2004.

[12] E. Telatar, "Capacity of multi-antenna Gaussian channels," *AT & T Bell Labs Internal Technical Memo*, June 1995.