Linear Space-Time Codes with Optimal Diversity-Multiplexing Tradeoff

Kambiz Azarian and J. Nicholas Laneman
Department of Electrical Engineering
University of Notre Dame
Notre Dame, IN 46556, USA
Email: {`kazarian,jnl`}`@nd.edu`

***Abstract*— In an effort to construct space-time codes (STC) that enjoy both optimality with respect to the diversity-multiplexing tradeoff (DMT) and encoding simplicity, we consider linear dispersion codes and establish a necessary and sufficient condition, in terms of the dispersion matrix, such that the resulting STC is DMT-optimal. We also present a simple construction that satisfies this condition, and is therefore DMT-optimal, for arbitrary number of transmit and receive antennas.**

I. INTRODUCTION

It is well known that multi-input multi-output (MIMO) wireless systems offer two types of enhancements, *i.e.*, throughput and reliability. It is also known that increased throughput comes at the price of reduced reliability, and vice versa, *i.e.*, that there exists a tradeoff between these two enhancements. In the context of frequency-flat Rayleigh block-fading channels, this tradeoff is known as the diversity-multiplexing tradeoff (DMT) originally characterized by Zheng and Tse [1].

Since the introduction of the DMT, there has been considerable research activity to construct space-time codes (STC) that trade throughput and reliability in an optimal way. For example, in [2] Yao and Wornell consider a 2×2 threaded STC, in which the diagonal and anti-diagonal threads are unitary transformations of QAM vectors. They show that through proper choice of the unitary transformations, the determinant of the codeword difference matrices can be bounded away from zero. They then use this result to prove the DMT-optimality of their construction. In [3], Dayal and Varanasi consider the same 2×2 threaded framework; however, they optimize the unitary transformations to get a better coding gain, as compared to the code of Yao and Wornell. Belfiore *et al.*[4] also utilize this framework and propose a 2×2 STC called the Golden code. The Golden code is DMT-optimal and exhibits very good performance in terms of codeword error probability.

Although all of the works mentioned above fall within the category of threaded STCs, these are not the only DMT-optimal codes in the literature. We mention only a few of these other codes here, as they use a different methodology than the focus of this work. We refer the interested reader to [5] for a more detailed survey. In [6], El Gemal *et al.* use lattice codes to construct DMT-optimal STCs, which they call LAST codes. In [5], Elia *et al.* use cyclic division algebras to construct explicit STCs that are DMT-optimal. In [7], Tavildar and Viswanath consider the problem of DMT-optimality under arbitrary fading distributions, i.e., approximately universal optimality. They show the existence of permutation codes that, if used in a parallel channel, are approximately optimal. They also show that these codes, if used in a D-BLAST fashion in a MIMO channel, provide DMT-optimality in the limit as the code length grows to infinity.

One distinguishing characteristic of the Yao-Wornell construction is its simplicity; however, this construction is limited to the case of two transmit and two receive antennas. In an effort to overcome this limitation, we consider a class of STCs called linear dispersion (LD) codes [8], and establish necessary and sufficient requirements, in terms of the the dispersion matrix, such that the resulting STC is DMT-optimal. In other words, we show that it is possible to use appropriate scalar codes to independently encode a number of streams, and then mix them using a proper dispersion matrix to obtain a DMT-optimal code. We do not address the stream encoding issue in this paper. We also present a simple construction that satisfies these requirements, and is therefore DMT-optimal, for arbitrary number of transmit and receive antennas.

The rest of this paper is organized as follows. In Section II, we detail the modeling assumptions and the LD space-time coding framework to be used. We also review in some detail the notion of diversity-multiplexing tradeoff, which is extensively used in the subsequent sections. In Section III, we state our main result, which is a necessary and sufficient condition for DMT-optimality of a LD STC. We further illustrate this result by an example. In Section IV, we present a simple construction for LD STCs and prove its DMT-optimality for arbitrary number of transit and receive antennas. Finally, in Section V, we offer some concluding remarks.

Before proceeding, we summarize some of the notations used throughout the paper. As usual, $\mathbb{C}$ is used to denote the set of complex numbers. Superscripts t and $*$ denote the transpose and conjugate transpose operations, respectively. $|A|$ denotes the cardinality of set A, and I_m denotes the $m \times m$ identity matrix. Finally, $\doteq$ denotes exponential equality, and $(x)^+$ denotes $\max\{x, 0\}$.

II. SYSTEM MODEL AND BACKGROUND

In this paper, we consider a MIMO wireless system with m transmit and n receive antennas. We denote the channel's degrees of freedom by $q \triangleq \min\{m, n\}$. We address a flat Rayleigh block-fading setup where the path gains remain constant over T consecutive symbol-intervals (*i.e.* a block), but change independently from one block to another. We further assume a coherent communication model resulting from the availability of channel state information (CSI) at the destination. Under these assumptions, the channel input-output relation is given by

$$\mathbf{y}_t = \sqrt{\rho}\tilde{\mathbf{H}}\mathbf{x}_t + \mathbf{n}_t \quad \text{for} \quad T \geq t \geq 1, \tag{1}$$

where $\mathbf{y}_t \in \mathbb{C}^n$, $\mathbf{x}_t \in \mathbb{C}^m$ and $\mathbf{n}_t \in \mathbb{C}^n$ represent the received signal, transmitted signal, and observed (unit-variance) additive white Gaussian noise, at time t, respectively. $\tilde{\mathbf{H}} \in \mathbb{C}^{n \times m}$ represents the channel matrix and ρ denotes the transmission power at each of the transmitting antennas.

Next, we define what we mean by a LD STC. Let $\{\mathbf{c}_i \in \mathbb{C}^T\}_{i=1}^q$ represent q *independently* encoded streams of symbols, each of rate R/q bits per symbol and length T. Also, define $\mathbf{c}, \mathbf{x}, \mathbf{y}$ and $\mathbf{H}$ as

$$\begin{aligned} \mathbf{c} &\triangleq [\mathbf{c}_1^t \cdots \mathbf{c}_q^t]^t, \\ \mathbf{x} &\triangleq [\mathbf{x}_1^t \cdots \mathbf{x}_T^t]^t, \\ \mathbf{y} &\triangleq [\mathbf{y}_1^t \cdots \mathbf{y}_T^t]^t \quad \text{and} \\ \mathbf{H} &\triangleq \text{diag}\{\tilde{\mathbf{H}}, \cdots, \tilde{\mathbf{H}}\}. \end{aligned}$$

We define an LD STC $\mathbf{x}$ as

$$\mathbf{x} \triangleq M\mathbf{c}, \tag{2}$$

where M is an $mT \times qT$ complex matrix called the dispersion matrix. We define matrices $\{M_{ti}\}_{t=1,i=1}^{T,q}$ (of dimension $m \times T$), such that $M = [M_{ti}]$. Let I be a nonempty subset of $\{1, \cdots, q\}$. The signature $\mathbf{x}_I$ of streams with indices $i \in I$ is given by

$$\mathbf{x}_I \triangleq M_I \mathbf{c}_I, \tag{3}$$

where $\mathbf{c}_I$ and M_I are derived from $\mathbf{c}$ and M, respectively, by deleting the entries corresponding to streams $i \notin I$. Matrices $\{M_{tI}\}_{t=1}^T$ (of dimension $m \times |I|T$), are defined such that $M_I = [M_{tI}]$.

Next we give a brief introduction to the Zheng and Tse's formulation of DMT [1]. This formulation assumes a family of space-time codes $\{\mathbf{x}_\rho\}$ indexed by the operating SNR ρ, such that the code $\mathbf{x}_\rho$ has rate $R(\rho)$, in bits per channel use (bpcu), and error probability $P_e(\rho)$. For this family, the multiplexing gain r and the diversity gain d are defined by

$$r \triangleq \lim_{\rho \to \infty} \frac{R(\rho)}{\log \rho} \quad \text{and} \quad d \triangleq -\lim_{\rho \to \infty} \frac{\log P_e(\rho)}{\log \rho}. \tag{4}$$

The optimal DMT yields the maximum possible diversity gain for every value of r, and is summarized in the following result.

Theorem 1 ([1]): The optimal diversity gain for the coherent block-fading MIMO channel with m transmit and n receive antennas, at multiplexing gain r, is given by $d(r) = f(r)$, where $f(\cdot)$ is the piecewise linear function joining the points $(k, (m-k)(n-k))$ for $k = 0, \ldots, q$. Moreover, there exists a code that achieves $d(r)$ for all block lengths $T \geq m + n - 1$.

III. A Necessary and Sufficient Condition for DMT-optimality of LD STCs

In this section we state our main result, which gives a necessary and sufficient condition for a LD STC to be DMT-optimal. It is very important to emphasize, however, that the condition is stated only in terms of the dispersion matrix M, *i.e.*, we do not address the problem of stream encoding. Put another way, M satisfying the condition ensures the existence of stream encoders such that the STC given by (2) achieves the optimal DMT. On the other hand, if the condition on M is not met, then there are no stream encoders such that the optimal DMT is achieved. We are now ready to state our result.

Theorem 2: A necessary and sufficient condition on M, for DMT-optimality of the LD STC given by (2), is that

$$\{\mathbf{u}_j^* M_{tI} | T \geq t \geq 1, |I| \geq j \geq 1\}, \tag{5}$$

be a set of $|I|T$ linearly independent vectors in $\mathbb{C}^{|I|T}$, with probability one, for any nonempty subset I of $\{1, \cdots, q\}$. In (5), $\{\mathbf{u}_j\}_{j=1}^m$ denotes a uniformly distributed set of orthogonal vectors in $\mathbb{C}^m$.

Proof: To prove the sufficiency part, we characterize $P_{e|H}$, *i.e.*, the joint ML error probability of the LD STC, averaged over the ensemble of Gaussian streams $\{\mathbf{c}_i\}$ and conditioned on a fixed channel realization $\mathbf{H} = H$. We then argue that for the average error probability P_e (averaged with respect to both the code and the channel ensembles) to be DMT-optimal, $P_{e|H}$ should be optimal with probability one.

To characterize $P_{e|H}$, we split the error event e into a number of disjoint error events $\{e_I\}$, where e_I denotes the event when only streams $\{\mathbf{c}_i\}_{i \in I}$ are decoded in error. We can then write

$$\begin{aligned} P_{e|H} &= \sum_I P_{e_I|H}, \quad \text{or} \\ d(r) &= \inf_I d_I(r), \end{aligned} \tag{6}$$

where $d(r)$ and $d_I(r)$ denote the diversity gains of P_e and P_{e_I}, respectively. Following [1], we upper bound the pairwise error probability $P_{e_I,p|H}$ by

$$P_{e_I,p|H} \leq \overset{-1}{\det}(I_{nT} + \frac{1}{2}\rho H \Sigma_{\mathbf{x}_I} H^*),$$

where $\Sigma_{\mathbf{x}_I}$ denotes the covariance matrix of $\mathbf{x}_I$. Using (3) and recalling that the streams are encoded using an ensemble of Gaussian codes, we get

$$\begin{aligned} P_{e_I,p|H} &\leq \overset{-1}{\det}(I_{mT} + \frac{1}{2}\rho M_I^* H^* H M_I), \\ &= \overset{-1}{\det}(I_{mT} + \frac{1}{2}\rho \sum_{j=1}^{q} \lambda_j \sum_{t=1}^{T} (\mathbf{u}_j^* M_{tI})^* (\mathbf{u}_j^* M_{tI})), \end{aligned}$$

where $\{\lambda_1 \geq \cdots \geq \lambda_q \geq 0\}$ and $\{\mathbf{u_j}\}_{j=1}^q$ are the eigenvalues and eigenvectors of $\tilde{H}^*\tilde{H}$, respectively. Since $\sum_{j=1}^{|I|} \lambda_j \sum_{t=1}^{T} (\mathbf{u}_j^* M_{tI})^* (\mathbf{u}_j^* M_{tI})$ and

$\sum_{j=|I|+1}^{q} \lambda_j \sum_{t=1}^{T}(\mathbf{u}_j^* M_{tI})^*(\mathbf{u}_j^* M_{tI})$ are both Hermitian and positive definite, we can write ([9], p. 367)

$$P_{e_I,p|H} \le \overset{-1}{\det}\Big(I_{mT} + \frac{1}{2}\rho \sum_{j=1}^{|I|} \lambda_j \sum_{t=1}^{T}(\mathbf{u}_j^* M_{tI})^*(\mathbf{u}_j^* M_{tI})\Big).$$

Now since $\{\mathbf{u}_j^* M_{tI}\}_{j=1,t=1}^{|I|,T}$ are linearly independent with probability one, we can use the Gram-Schmidt orthonormalization process [9] to show that, with probability one, $\sum_{j=1}^{|I|} \lambda_j \sum_{t=1}^{T}(\mathbf{u}_j^* M_{tI})^*(\mathbf{u}_j^* M_{tI})$ has exactly T eigenvalues corresponding to every λ_j for $|I| \ge j \ge 1$, and of the same exponential order. Thus

$$P_{e_I,p|H} \dot{\le} \prod_{j=1}^{|I|}(1+\rho\lambda_j)^{-T},$$

which leads to an outage event [1]

$$O_I = \{\alpha | \sum_{j=1}^{|I|}(1-\alpha_j)^+ \le \frac{|I|}{q} R\}, \tag{7}$$

where $\lambda_j \dot{=} \rho^{-\alpha_j}$. Recalling that the outage set for a standard $m \times n$ MIMO is [1]

$$O_{m\times n} = \{\alpha | \sum_{j=1}^{q}(1-\alpha_j)^+ \le R\},$$

we realize that for all $I \ne \{1,\cdots,q\}$, $O_I \subset O_{m\times n}$ and thus

$$d_I(r) \ge d_{m\times n}(r), \forall I \ne \{1,\cdots q\}. \tag{8}$$

In this expression, $d_{m\times n}(r)$ denotes the optimal diversity gain for an $m \times n$ MIMO channel. For $I = \{1,\cdots,q\}$, $O_I = O_{m\times n}$ and thus $d_I(r) = d_{m\times n}(r)$. This, together with (8) and (6), gives $d(r) = d_{m\times n}(r)$, which concludes the proof for the sufficiency part.

Due to space limitations, we do not give the proof for the necessity part, though the main idea is the same, *i.e.*, if (5) does not hold with probability one, then the signature of certain groups of streams, as a whole, misses the best eigenvalue of the channel with nonzero probability. This clearly prevents the LD STC from achieving the optimal DMT.

Before illustrating the application of Theorem 2, we notice that working with an ensemble of random stream encoders, rather than a specific set of them, has relieved us from dealing with codeword difference matrices and has resulted in a concise description of the conditions to be met by the dispersion matrix such that the resulting STC is DMT-optimal. This, however, comes at the price of a weaker statement, *i.e.*, Theorem 2, only asserts the *existence* of good scalar codes that, when used in conjunction with an appropriate dispersion matrix, result in a DMT-optimal STC.

Next, we illustrate the application of Theorem 2 by examining the 2×2 construction of Yao and Wornell [2]. In this construction, the matrix M is set to

$$M = \begin{bmatrix} \cos\theta_1 & -\sin\theta_1 & 0 & 0 \\ 0 & 0 & \cos\theta_2 & -\sin\theta_2 \\ 0 & 0 & \sin\theta_2 & \cos\theta_2 \\ \sin\theta_1 & \cos\theta_1 & 0 & 0 \end{bmatrix}. \tag{9}$$

Now for $I = \{1\}$,

$$\begin{aligned} \mathbf{u}_1^* M_{11} &= \begin{bmatrix} u_1(1)\cos\theta_1 & -u_1(1)\sin\theta_1 \end{bmatrix}, \\ \mathbf{u}_1^* M_{21} &= \begin{bmatrix} u_1(2)\sin\theta_1 & u_1(2)\cos\theta_1 \end{bmatrix}, \end{aligned}$$

which are obviously linearly independent with probability one (if $u_1(1) = 0$ or $u_1(2) = 0$, then one of the vectors vanishes). The case of $I = \{2\}$ is very similar to that of $I = \{1\}$. For $I = \{1,2\}$,

$$\mathbf{u}_1^* M_{1\{1,2\}} = \begin{bmatrix} u_1(1)\cos\theta_1 \\ -u_1(1)\sin\theta_1 \\ u_1(2)\cos\theta_2 \\ -u_1(2)\sin\theta_2 \end{bmatrix}^t \quad \mathbf{u}_2^* M_{1\{1,2\}} = \begin{bmatrix} u_2(1)\cos\theta_1 \\ -u_2(1)\sin\theta_1 \\ u_2(2)\cos\theta_2 \\ -u_2(2)\sin\theta_2 \end{bmatrix}^t$$

and

$$\mathbf{u}_1^* M_{2\{1,2\}} = \begin{bmatrix} u_1(2)\sin\theta_1 \\ u_1(2)\cos\theta_1 \\ u_1(1)\sin\theta_2 \\ u_1(1)\cos\theta_2 \end{bmatrix}^t \quad \mathbf{u}_2^* M_{2\{1,2\}} = \begin{bmatrix} u_2(2)\sin\theta_1 \\ u_2(2)\cos\theta_1 \\ u_2(1)\sin\theta_2 \\ u_2(1)\cos\theta_2 \end{bmatrix}^t$$

which are again seen to be linearly independent with probability one. Thus the construction is DMT-optimal.

IV. A CONSTRUCTION FOR DMT-OPTIMAL LD STC

In this section, we present an example construction for an LD STC with arbitrary number of transmit and receive antennas, and further prove its DMT-optimality through the use of Theorem 2. To motivate the construction, notice that Theorem 2 requires $\{\mathbf{u}_j^* M_{tI} | T \ge t \ge 1, |I| \ge j \ge 1\}$ to be a set of $|I|T$ linearly independent vectors, for most choices of $\{\mathbf{u}_j\}$. Let $\mathbf{u}_j$ be a unit vector with only one nonzero element (one) at position j. Also, let $I = \{1\}$. With these choices, $\{\mathbf{u}_j^* M_{tI} | T \ge t \ge 1, |I| \ge j \ge 1\}$ reduces to the collection of the first rows of M_{t1}, for $T \ge t \ge 1$. Now, one way to make sure that this collection is linearly independent is by setting them to an (arbitrary) orthogonal set $\{\mathbf{v}_j\}$ (as we mentioned in the previous section, for this particular choice of $\mathbf{u}_1$, $\{\mathbf{u}_1 M_{21}\}$ in the Yao-Wornell construction vanishes to zero, however this does not hurt the DMT-optimality of the construction as the probability of such $\mathbf{u}_1$ is zero). This same argument could be made for the collection of the kth ($m \ge k \ge 1$) rows of M_{ti} by modifying $\mathbf{u}_1$ to have its nonzero element at the kth position, and by setting $I = \{i\}$. With these ideas in mind, we next detail our construction.

The construction is best illustrated through an example. Set $m = q = T = 3$, and let $\{\mathbf{v}_j\}_{j=1}^3$, $\{\tilde{\mathbf{v}}_j\}_{j=1}^3$ and $\{\bar{\mathbf{v}}_j\}_{j=1}^3$ be three sets of orthogonal vectors in $\mathbb{C}^{1\times 3}$. Also, let $\{\phi_j\}_{j=1}^3$ be three *distinct* complex numbers. Now, the desired STC is constructed by setting $\{M_{t,\{1,2,3\}}\}_{t=1}^{3}$ to

$$M_{1,\{1,2,3\}} = \begin{bmatrix} \phi_1\mathbf{v}_1 & \phi_2\tilde{\mathbf{v}}_1 & \phi_3\bar{\mathbf{v}}_1 \\ \phi_1\mathbf{v}_2 & \phi_2\tilde{\mathbf{v}}_2 & \phi_3\bar{\mathbf{v}}_2 \\ \phi_1\mathbf{v}_3 & \phi_2\tilde{\mathbf{v}}_3 & \phi_3\bar{\mathbf{v}}_3 \end{bmatrix}, \tag{10}$$

$$M_{2,\{1,2,3\}} = \begin{bmatrix} \phi_2\mathbf{v}_3 & \phi_3\tilde{\mathbf{v}}_3 & \phi_1\bar{\mathbf{v}}_3 \\ \phi_2\mathbf{v}_1 & \phi_3\tilde{\mathbf{v}}_1 & \phi_1\bar{\mathbf{v}}_1 \\ \phi_2\mathbf{v}_2 & \phi_3\tilde{\mathbf{v}}_2 & \phi_1\bar{\mathbf{v}}_2 \end{bmatrix}, \tag{11}$$

$$M_{3,\{1,2,3\}} = \begin{bmatrix} \phi_3\mathbf{v}_2 & \phi_1\tilde{\mathbf{v}}_2 & \phi_2\bar{\mathbf{v}}_2 \\ \phi_3\mathbf{v}_3 & \phi_1\tilde{\mathbf{v}}_3 & \phi_2\bar{\mathbf{v}}_3 \\ \phi_3\mathbf{v}_1 & \phi_1\tilde{\mathbf{v}}_1 & \phi_2\bar{\mathbf{v}}_1 \end{bmatrix}. \tag{12}$$

It is important to realize that the only requirement on $\{\mathbf{v}_j\}$, $\{\tilde{\mathbf{v}}_j\}$ and $\{\bar{\mathbf{v}}_j\}$, as far as DMT-optimality is concerned, is that they are sets of orthogonal vectors, *e.g.*, we could have chosen $\{\mathbf{v}_j\}$, $\{\tilde{\mathbf{v}}_j\}$ and $\{\bar{\mathbf{v}}_j\}$ to be identical, or for that matter any permutations of one another. Also, while we have chosen $M_{<t>_3+1,i}$ to be (up to a scaling factor) the circular shift down of $M_{t,i}$, this need not be the case, *e.g.*, we could have set $M_{<t>_3+1,2}$ to be a circular shift up of $M_{t,2}$.

To get an appreciation of why the construction achieves the optimal DMT, let $\mathbf{v}_j = \tilde{\mathbf{v}}_j = \bar{\mathbf{v}}_j$ be the unit vector with a single nonzero element at the jth position. It can then be seen that, *e.g.*, transmit antenna one, two and three transmit $\phi_1 c_1(1) + \phi_2 c_2(1) + \phi_3 c_3(1)$, $\phi_2 c_1(1) + \phi_3 c_2(1) + \phi_1 c_3(1)$, and $\phi_3 c_1(1) + \phi_1 c_2(1) + \phi_2 c_3(1)$ over symbol-interval $t = 1$, 2 and 3, respectively ($c_i(t)$ is the tth symbol in stream $\mathbf{c}_i$). Thus every symbol in each of the streams is transmitted by all of the antennas. This gives the intuition why the construction achieves the maximal diversity gain of the channel. To understand why the construction achieves the maximal multiplexing gain, notice that on the average, three new symbols are introduced per symbol-interval, which is equal to the channel's degrees of freedom. The following theorem establishes the DMT-optimality of the construction formally.

Theorem 3: The construction illustrated by (10), (11) and (12) is DMT-optimal.

Proof: Using Theorem 2, to establish the DMT-optimality of the construction, we need to show that $\{\mathbf{u}_j^* M_{tI} | m \geq t \geq 1, |I| \geq j \geq 1\}$ is a set of $m|I|$ linearly independent vectors with probability one (without loss of generality, we set $I = \{1, \cdots, |I|\}$). Toward this end, we show that if

$$\sum_{t=1}^{m}\sum_{j=1}^{|I|} a_{t,j}\mathbf{u}_j^* M_{t,I} = 0, \tag{13}$$

then $a_{t,j} = 0, \forall t, j$ with probability one. First we notice that (13) is equivalent to

$$\sum_{t=1}^{m}\sum_{j=1}^{|I|} a_{t,j}\mathbf{u}_j^* M_{t,i} = 0, \quad |I| \geq i \geq 1. \tag{14}$$

However, since the corresponding rows of $\{M_{t,i}\}_{t=1}^{m}$ are linearly independent, for every value of i, equation (14) can be spilt-up into m equations. Let us denote the vector derived by element-wise conjugation of $\mathbf{u}_j$ by $\hat{\mathbf{u}}_j$ (note that $\hat{\mathbf{u}}_j \in \mathbb{C}^{m\times 1}$, not $\mathbb{C}^{1\times m}$). Also, let us denote the vector derived by circularly down-shifting vector $\mathbf{u}_j$ for k times by $\pi(\mathbf{u}_j, k)$. Using these notations, and through inspection, we can write the set of equations resulting from (14) as

$$\sum_{t=1}^{m}\sum_{j=1}^{|I|} a_{t,j}\phi_{<t+i-2>_m+1}\pi(\hat{\mathbf{u}}_{\mathbf{j}}, t-1) = \mathbf{0}, \quad |I| \geq i \geq 1,$$

which can in turn be written as $\mathbf{U}_I\mathbf{a}_I = \mathbf{0}$, where $\mathbf{a}_I \triangleq [a_{1,1}, a_{2,1}, \cdots, a_{m,|I|}]^t$ and

$$\mathbf{U}_I \triangleq \begin{bmatrix} \phi_1\hat{\mathbf{u}}_1 & \phi_2\pi(\hat{\mathbf{u}}_1, 1) & \cdots & \phi_m\pi(\hat{\mathbf{u}}_{|I|}, m-1) \\ \phi_2\hat{\mathbf{u}}_1 & \phi_3\pi(\hat{\mathbf{u}}_1, 1) & \cdots & \phi_1\pi(\hat{\mathbf{u}}_{|I|}, m-1) \\ \vdots & \vdots & & \vdots \\ \phi_{|I|}\hat{\mathbf{u}}_1 & \phi_{|I|+1}\pi(\hat{\mathbf{u}}_{\mathbf{1}}, 1) & \cdots & \phi_{|I|-1}\pi(\hat{\mathbf{u}}_{|I|}, m-1) \end{bmatrix}.$$

Now, for $\mathbf{a}_I$ to be $\mathbf{0}$, $\mathbf{U}_I$ needs to be full rank. Given that $\{\phi_j\}_{j=1}^{m}$ are distinct, one can show that if

$$\mathbf{U}_j \triangleq \begin{bmatrix} \hat{\mathbf{u}}_j & \pi(\hat{\mathbf{u}}_j, 1) & \cdots & \pi(\hat{\mathbf{u}}_j, m-1) \end{bmatrix}, \quad \forall |I| \geq j \geq 1$$

are full rank, then $\mathbf{U}_I$ is also full rank. Note, however, that $\{\mathbf{U}_j\}$ is a set of circulant matrices. The eigenvalues of a circulant matrix are given by the discrete fourier transform (DFT) of its first row, *i.e.*, DFT of $\mathbf{u}_j$. Since $\{\mathbf{u}_j\}$ are uniformly distributed, $\{\mathbf{U}_j\}$ are full rank with probability one. This concludes the proof.

V. CONCLUSION

We considered LD codes and established a necessary and sufficient condition on the dispersion matrix, such that the resulting STC is DMT-optimal. In other words, we showed that it is possible to use appropriate scalar codes to independently encode a number of streams and then mix them using an appropriate dispersion matrix to get a DMT-optimal STC. We also presented a simple construction that satisfies these requirements, and is thus DMT-optimal, for arbitrary number of transmit and receive antennas.

REFERENCES

[1] L. Zheng and D. N. C. Tse, "Diversity and Multiplexing: A Fundamental Tradeoff in Multiple Antenna Channels," *IEEE Trans. Info. Theory*, 49:1073-1096, May 2003.

[2] H. Yao, G. W. Wornell, "Structured Space-time Block Codes with Optimal Diversity-Multiplexing Tradeoff and Minimum Delay," *IEEE Globecom 03*, 4:1941-1945, Dec. 2003.

[3] P. Dayal and M. K. Varanasi, "An Optimal Two Transmit Antenna Space Time Code and Its Stacked Extension," *Proc. Asilomar Conf. Signals, Systems and Computers*, M onterey, CA, Nov. 2003.

[4] J. C. Belfiore, G. Rekaya, and E. Viterbo, "The Golden Code: A 2 × 2 Full-rate Space-time Code with Nonvanishing Determinants," IEEE Trans. Inf. Theory, 51:1432, Apr. 2005.

[5] P. Elia, K. Raj Kumar, S. A. Pawar, P. Vijay Kumar and Hsiao-Feng Lu, "Explicit Space-Time Codes Achieving the Diversity-Multiplexing Gain Tradeof," *IEEE Trans. Info. Theory*, 52:3869-3884, Sept. 2006.

[6] H. El Gamal, G. Caire and M. O. Damen, "Lattice Coding and Decoding Achieve the Optimal DiversityMultilpexing Tradeoff of MIMO Channels," *IEEE Trans. Inf. Theory*, 50:968, Jun. 2004.

[7] S. Tavildar and P. Viswanath, "Approximately Universal Codes over Slow-Fading Channels," *IEEE Trans. Inf. Theory*, 52:3233 Jul., 2006.

[8] B. Hassibi and B. M. Hochwald, "High-Rate Codes that are Linear in Space and Time," *IEEE Trans. Info. Theory*, 48:1804, Jul. 2002.

[9] R. A. Horn and C. R. Johnson, "Matrix Analysis," *Cambridge University Press*, 1999.