Linear Toeplitz Space Time Block Codes

Jian-Kang Zhang, Jing Liu and Kon Max Wong
Department of Electrical and Computer Engineering, McMaster University, Hamilton, Ontario, Canada.
Email: (jkzhang, wong)@mail.ece.mcmaster.ca; jliu@grads.ece.mcmaster.ca

***Abstract*— In this paper we consider coherent flat fading wireless communication systems with multiple transmitter antennas and single receiver antenna (MISO). We propose a Toeplitz linear space time block code (STBC) that converts an original MISO flat fading channel into a Toeplitz virtual multiple inputs multiple outputs (MIMO) channel. We show that our proposed code has the following main features: (a) The symbol transmission rate is $\frac{T-M+1}{T}$, where M is the number of transmitter antenna and T is the number of channel uses ($T > M$). (b) Linear receivers (zero-forcing and minimum mean square error) can extract full diversity. Moreover, when the channel coefficients are independent and the maximum likelihood (ML) detector is employed, our Toeplitz STBC minimizes the exact worst case average pair-wise error probability. (c) When channels are correlated, we design our Toeplitz STBC that minimizes the exact average worst case pair-wise error probability. By transforming this non-convex optimization problem into a convex one, the problem can be solved efficiently by employing an interior point method. In particular, when the design criterion in question is approximated by the Chernoff bound, we obtain a closed form solution. (d) Finally, for the independent MISO flat fading system, we prove that our proposed codes can approach the optimal diversity-vs-multiplexing tradeoff developed by Zheng and Tse with a linear zero-forcing receiver when the number of channel uses is large.**

I. Introduction

In this paper, we consider a coherent flat fading wireless communication system with multiple transmitter and the single receiver antenna; i.e., a multiple inputs single output (MISO) system, which can be represented in a compact vector form as

$$r = \mathbf{x}^T\mathbf{h} + \xi, \tag{1}$$

where r is a received signal, $\mathbf{h}$ is an $M\times 1$ channel vector, $\mathbf{x}$ is an $M\times 1$ transmitting signal vector and ξ is a complex noise. Throughout this paper, we adopt the following assumptions: (a) The channel $\mathbf{h}$ is circularly-symmetric complex Gaussian distributed, with zero-mean, and positive definite covariance matrix $\mathbf{\Sigma}$; (b) ξ is a circularly-symmetric complex Gaussian noise with variance σ^2. Our goal is to design linear space-time block codes that minimize the worst case average pair-wise error probability and asymptotically achieve the optimal diversity-multiplexing tradeoff [1]. Utilizing channel covariance information, the optimal transmitter design has been pursued on the basis of a capacity criterion [2]–[6]. For MISO communication systems, Zhou and Giannakis [7] designed the precoder that minimizes the upper bound of the average symbol error probability (SEP) based on maximum ratio combining receiver and orthogonal space-time codes [8]–[10]. However, orthogonal codes suffer from a limited transmission rate [11]–[14], and thus do not achieve full capacity in MIMO channels [15]. To overcome this, Hassibi and Hochwald [16] proposed linear dispersion codes. Recent research [17]–[19] based on number theory has shown that it is possible to design linear space-time block codes and dispersion codes [20] which are full rate and full diversity without information loss. The main issue on these current designs is that the coding gain vanishes rapidly as the constellation size increases. Therefore, full rate full diversity non-vanishing space-time code designs have recently drawn much attention [21]–[27] due to an important potential [21] that such structured space-time codes could achieve the optimal diversity-vs-multiplexing tradeoff developed by Zheng and Tse [1]. More recently, a class of random lattice space-time codes [28] has been designed to achieve the optimal diversity-multiplexing tradeoff [1] for a general coherent MIMO flat-fading channels under generalized minimum Euclidean distance lattice decoding by exploiting Erez and Zamir's nested lattice scheme [29]. Motivated by [1], [21], [28], [30], we propose Toeplitz space-time codes to asymptotically achieve the optimal diversity-multiplexing tradeoff [1] for MISO channels with a linear ZF receiver.

Notation: Matrices and column vectors are denoted by uppercase boldface characters and lowercase boldface characters, respectively. The (i, j)-th entry of $\mathbf{B}$ is denoted by $[\mathbf{B}]_{i,j}$. Notation $\mathbf{I}_K$ denotes a $K\times K$ identity matrix. $\mathbf{B}^T$ the transpose, and $\mathbf{B}^H$ the conjugate transpose.

II. Toeplitz space-time block codes

First we introduce the definition of Toeplitz space-time block codes.

Definition 1: Let $\boldsymbol{\alpha} = [\alpha_1, \alpha_2, \cdots, \alpha_L]^T$. Then, a $(K + L - 1)\times K$ Toeplitz matrix generated by $\boldsymbol{\alpha}$ and a positive integer K, denoted by $\mathcal{T}(\boldsymbol{\alpha}, L, K)$, is defined as

$$[\mathcal{T}(\boldsymbol{\alpha}, L, K)]_{i,j} = \begin{cases} \alpha_{i-j+1}, & \text{if } i \geq j \text{ and } i-j < L \\ 0, & \text{otherwise.} \end{cases} \tag{2}$$

which can be explicitly written out as

$$\mathcal{T}(\boldsymbol{\alpha}, L, K) = \begin{pmatrix} \alpha_1 & 0 & \dots & 0 \\ \alpha_2 & \alpha_1 & \dots & 0 \\ \vdots & \alpha_2 & \ddots & \vdots \\ \alpha_L & \ddots & \ddots & \alpha_1 \\ 0 & \ddots & \ddots & \alpha_2 \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \ddots & 0 & \alpha_L \end{pmatrix}_{(K+L-1)\times K}. \tag{3}$$

Then, a Toeplitz space-time block code $\mathcal{X}_{\mathbf{A}}(\mathbf{s})$ is defined as

$$\mathcal{X}_{\mathbf{A}}(\mathbf{s}) = \mathcal{T}(\mathbf{s}, N, M)\mathbf{A} \tag{4}$$

where $\mathbf{A}$ is an $M \times M$ invertible matrix. ■

At time slot t, the tth row of $\mathcal{X}_{\mathbf{A}}(\mathbf{s})$ is fed to M transmitter antennas for transmission. At the receiver, all the T received signals can be written as a vector such that

$$\mathbf{r} = \mathcal{X}_{\mathbf{A}}(\mathbf{s})\mathbf{h} + \boldsymbol{\xi}, \tag{5}$$

where $\mathbf{r} = [r_1, r_2, \cdots, r_T]^T$ and $\boldsymbol{\xi} = [\xi_1, \xi_2, \cdots, \xi_T]^T$. Substituting (4) into (5), we have

$$\mathbf{r} = \mathcal{T}(\widetilde{\mathbf{h}}, M, N)\,\mathbf{s} + \boldsymbol{\xi}, \tag{6}$$

where $N = T - M + 1$, $\widetilde{\mathbf{h}} = \mathbf{A}\mathbf{h}$ and we have utilized the fact that

$$\mathcal{X}_{\mathbf{A}}(\mathbf{s})\mathbf{h} = \mathcal{T}(\widetilde{\mathbf{h}}, M, N)\,\mathbf{s}. \tag{7}$$

Now in (6), the original MISO channel is transformed into a Toeplitz visual MIMO channel. Such a channel is a special convolutive channel and hence, we can utilize the efficient Vitterbi algorithm [31] to detect the signal $\mathbf{s}$ with perfect channel knowledge at the receiver. Also, we can take advantage of the second order statistics to blindly identify the channel. In addition, we see from Definition 1 that the symbol transmission rate of our Toeplitz space-time codes is $R = \frac{T-M+1}{T}$ per channel use. Therefore, for a fixed M, the transmission rate R can approach one if channel uses are sufficiently large.

III. DIVERSITY OF ZF RECEIVER FOR TOEPLITZ STBCS

In this section, we will show that our Toeplitz space-time block codes can provide full diversity even for the linear ZF receiver. To do that, we first establish the following lemma.

Lemma 1: There exists a positive definite constant C such that for any nonzero vector $\boldsymbol{\alpha}$, the following inequality holds,

$$C\|\boldsymbol{\alpha}\|^{2K} \le \det\left(\mathcal{T}^H(\boldsymbol{\alpha}, L, K)\mathcal{T}(\boldsymbol{\alpha}, L, K)\right) \le \|\boldsymbol{\alpha}\|^{2K}. \tag{8}$$

■

Proof: First we notice that the diagonal entries of matrix $\mathcal{T}^H(\boldsymbol{\alpha}, L, K)\mathcal{T}(\boldsymbol{\alpha}, L, K)$ are same and equal to

$$\left[\mathcal{T}^H(\boldsymbol{\alpha}, L, K)^H\mathcal{T}(\boldsymbol{\alpha}, L, K)\right]_{k,k} = \|\boldsymbol{\alpha}\|^2.$$

By employing Hardamard's inequality [32], the right side of inequality (8) can be obtained immediately. The left hand side is proved as follows. For $\boldsymbol{\alpha} \neq \mathbf{0}$, let $\gamma_k = |\alpha_k|/\|\boldsymbol{\alpha}\|$ and $\alpha_k = |\alpha_k|\exp(j2\pi\theta_k)$. Then, we have

$$\det\left(\mathcal{T}^H(\boldsymbol{\alpha}, L, K)^H\mathcal{T}(\boldsymbol{\alpha}, L, K)\right) = \|\boldsymbol{\alpha}\|^{2N}D(\boldsymbol{\gamma}, \boldsymbol{\theta}), \tag{9}$$

where $\boldsymbol{\gamma} = [\gamma_1, \gamma_2, \cdots, \gamma_L]^T$ and $\boldsymbol{\theta} = [\theta_1, \theta_2, \cdots, \theta_L]^T$. Let $\mathbb{D}$ denote the feasible set of $D(\boldsymbol{\gamma}, \boldsymbol{\theta})$; i.e.,

$$\mathcal{D} = \{(\boldsymbol{\gamma}, \boldsymbol{\theta}) : \|\boldsymbol{\gamma}\|^2{=}1, 0 \le \theta_\ell \le 2\pi \text{ for } \ell = 1, 2, \cdots, L\}$$

Since $D(\boldsymbol{\gamma}, \boldsymbol{\theta})$ is polynomial with respect to $\gamma_\ell, \cos\theta_\ell$ and $\sin\theta_\ell$ for $\ell = 1, 2, \cdots, L$, there exists the minimum value of $D(\boldsymbol{\gamma}, \boldsymbol{\theta})$ in the feasible set $\mathbb{D}$. Let $C = \min_{(\boldsymbol{\gamma}, \boldsymbol{\theta}) \in \mathbb{D}} D(\boldsymbol{\gamma}, \boldsymbol{\theta})$, and it suffices to prove that $C > 0$. Since $\boldsymbol{\alpha} \neq \mathbf{0}$, without loss of generality, we can always assume that $\alpha_1 \neq 0$. Otherwise, we can permute the rows of $\mathcal{T}(\boldsymbol{\alpha}, L, K)$ such that the first entry is nonzero. In this case, $\mathcal{T}(\boldsymbol{\alpha}, L, K)$ defined by (3) can be partioned as

$$\mathcal{T}(\boldsymbol{\alpha}, L, K) = \begin{pmatrix} \mathbf{B} \\ \mathbf{C} \end{pmatrix},$$

where $\mathbf{B}$ contains the first L rows of $\mathcal{T}(\boldsymbol{\alpha}, L, K)$ and is therefore lower triangular, and $\mathbf{C}$ denotes the remaining submatrix of $\mathcal{T}(\boldsymbol{\alpha}, L, K)$. Hence, we can write $\mathcal{T}^H(\boldsymbol{\alpha}, L, K)\mathcal{T}(\boldsymbol{\alpha}, L, K) = \mathbf{B}^H\mathbf{B} + \mathbf{C}^H\mathbf{C}$ and as a result, $\det\left(\mathcal{T}^H(\boldsymbol{\alpha}, L, K)\mathcal{T}(\boldsymbol{\alpha}, L, K)\right) \ge \det(\mathbf{B}^H\mathbf{B}) + \det(\mathbf{C}^H\mathbf{C}) \ge \det(\mathbf{B}^H\mathbf{B}) = |\alpha_1|^{2K} > 0$. Therefore, C is positive. □

By Lemma 1, it is immediate to get

Corollary 1: The positive constant C in Lemma 1 renders the following inequality $\left[\left(\mathcal{T}^H(\boldsymbol{\alpha}, L, K)\mathcal{T}(\boldsymbol{\alpha}, L, K)\right)^{-1}\right]_{k,k}^{-1} \ge C\|\boldsymbol{\alpha}\|^2$ is true for $k = 1, 2, \cdots, K$. ■

We are now in the position to formally state the first of our main results.

Theorem 1: The Toeplitz space-time block code provides full diversity for the zero-forcing receiver when D-ary PAM, PSK or square QAM signals are transmitted. ■

Proof: First we need to derive the expressions of symbol error probabilities for D-ary PAM, PSK and square QAM modulations when the ZF receiver is employed. for notional simplicity, let $\mathbf{P} = \mathcal{T}^H(\widetilde{\mathbf{h}}, M, N)\mathcal{T}(\widetilde{\mathbf{h}}, M, N)$.

1) PAM signals: The SEP of the ZF receiver for D-ary PAM signal s_k is

$$\mathrm{P}_{\mathrm{PAM}}(\mathbf{h}, s_k) = \frac{2(D-1)}{D}Q\left(\sqrt{\frac{3E_s}{(D^2-1)\sigma^2[\mathbf{P}^{-1}]_{k,k}}}\right), \tag{10}$$

where D is the constellation size and $Q(z) = \frac{1}{\sqrt{\pi}}\int_z^\infty e^{-x^2}dx$. Further, (10) is upper bounded by

$$\mathrm{P}_{\mathrm{PAM}}(\mathbf{h}, s_k) \le \frac{D-1}{D}\exp\left(-\frac{3E_s}{(D^2-1)\sigma^2[\mathbf{P}^{-1}]_{k,k}}\right)$$

2) PSK signals: The SEP of a ZF receiver for the PSK signal s_k is given by

$$\mathrm{P}_{\mathrm{PSK}}(\mathbf{h}, s_k) = \frac{1}{\pi}\int_0^{(D-1)\pi/D}\exp\left(-\frac{E_s\sin^2(\pi/D)}{2\sigma^2[\mathbf{P}^{-1}]_{k,k}\sin^2\theta}\right)d\theta,$$

which can be upper bounded by

$$\mathrm{P}_{\mathrm{PSK}}(\mathbf{h}, s_k) \le \frac{(D-1)}{D}\exp\left(-\frac{E_s\sin^2(\pi/D)}{2\sigma^2[\mathbf{P}^{-1}]_{k,k}}\right). \tag{11}$$

3) Square QAM signals: The SEP of a ZF receiver for the square QAM signal s_k is given by

$$\begin{aligned}\mathrm{P}_{\mathrm{QAM}}(\mathbf{h}, s_k) = {} & 4\left(1-\frac{1}{\sqrt{D}}\right)Q\left(\sqrt{\frac{3E_s}{2(D-1)\sigma^2[\mathbf{P}^{-1}]_{k,k}}}\right) \\ & - 4\left(1-\frac{1}{\sqrt{D}}\right)^2 Q^2\left(\sqrt{\frac{3E_s}{2(D-1)\sigma^2[\mathbf{P}^{-1}]_{k,k}}}\right).\end{aligned} \tag{12}$$

It is convenient to use the following alternative expressions for the Q function and the Q^2 function [33]

$$Q(z) = \frac{1}{\pi}\int_0^{\pi/2}\exp\left(-\frac{z^2}{2\sin^2\theta}\right)d\theta \tag{13}$$

$$Q^2(z) \quad = \quad \frac{1}{\pi}\int_0^{\pi/4} \exp\left(-\frac{z^2}{2\sin^2\theta}\right)\, d\theta \qquad (14)$$

Substituting (13) and (14) into (12), we obtain

$$\begin{aligned}
& \mathrm{P}_{\mathrm{QAM}}(\mathbf{h}, s_k) \\
= & \ \frac{4}{\pi}\left(1-\frac{1}{\sqrt{D}}\right)\int_0^{\pi/4} \exp\left(-\frac{3E_s}{2(D-1)\sigma^2[\mathbf{P}^{-1}]_{k,k}\sin^2\theta}\right) d\theta \\
+ & \ \frac{4}{\pi\sqrt{D}}\left(1-\frac{1}{\sqrt{D}}\right)\int_{\pi/4}^{\pi/2} \exp\left(-\frac{3E_s}{2(D-1)\sigma^2[\mathbf{P}^{-1}]_{k,k}\sin^2\theta}\right) d\theta.
\end{aligned}$$

Similarly, this can be upper bounded by

$$\mathrm{P}_{\mathrm{QAM}}(\mathbf{h}, s_k) \leq \frac{D-1}{D}\exp\left(-\frac{3E_s}{2(D-1)\sigma^2[\mathbf{P}^{-1}]_{k,k}}\right).$$

Therefore, we obtain a unified upper bound on the symbol error probability for D-ary PAM, PSK and QAM signals; i.e.,

$$\mathrm{P}(\mathbf{h}, s_k) \leq \frac{D-1}{D}\exp\left(-\frac{a}{[\mathbf{P}^{-1}]_{k,k}}\right), \qquad (15)$$

where $a_{\mathrm{PAM}} = 3E_s/(D^2-1)\sigma^2, a_{\mathrm{PSK}} = \frac{E_s\sin^2(\pi/D)}{2\sigma^2}$ and $a_{\mathrm{QAM}} = 3E_s/2(D-1)\sigma^2$. Combining (15) with Lemma 1, we have that the arithmetic mean of all SEPs is upper bounded by

$$\mathrm{P}^{(\mathrm{ZF})}(\mathbf{h}) \leq \frac{D-1}{D}\exp\left(-aC\|\mathbf{h}\|^2\right) \qquad (16)$$

Therefore, taking an average over the random vector $\mathbf{h}$ yields

$$\mathrm{E}\left[\mathrm{P}(\mathbf{h})\right] \leq \frac{D-1}{D}\det(\mathbf{I}+aC\mathbf{\Sigma})^{-1} \leq \frac{D-1}{D}\det(C\mathbf{\Sigma})^{-1}a^{-M}. \quad (17)$$

This completes the proof of Theorem 1. □

IV. Design of optimal Toeplitz STBCs

In this section, we will design the matrix $\mathbf{A}$ in a Toeplitz space-time block code such that the worst case pair-wise error probability is minimized when a maximum likelihood detector is employed.

Given a channel realization $\mathbf{h}$, the probability $P(\mathbf{s} \rightarrow \mathbf{s}'|\mathbf{h})$ of transmitting $\mathbf{s}$ and deciding in favor of $\mathbf{s}' \neq \mathbf{s}$ with the ML detector is given by [34]

$$P(\mathbf{s} \rightarrow \mathbf{s}'|\mathbf{h}) = Q\left(\frac{d(\mathbf{s}, \mathbf{s}')}{2\sigma}\right), \qquad (18)$$

where $d(\mathbf{s}, \mathbf{s}')$ is the Euclidean distance between $\mathcal{T}(\widetilde{\mathbf{h}}, M, N)\mathbf{s}$ and $\mathcal{T}(\widetilde{\mathbf{h}}, M, N)\mathbf{s}'$, $d^2(\mathbf{s}, \mathbf{s}') = (\mathbf{s} - \mathbf{s}')^H\mathcal{T}^H(\widetilde{\mathbf{h}}, M, N)\mathcal{T}(\widetilde{\mathbf{h}}, M, N)(\mathbf{s} - \mathbf{s}')$. From (7) we have $d^2(\mathbf{s}, \mathbf{s}') = \mathbf{h}^H\mathcal{X}_{\mathbf{A}}^H(\mathbf{e})\mathcal{X}_{\mathbf{A}}(\mathbf{e})\,\mathbf{h}$, where $\rceil = \mathbf{s} - \mathbf{s}'$. By employing (13) and taking the average of (18) over the random vector $\mathbf{h}$, the average pair-wise error probability can be written as

$$P_{\mathbf{A}}(\mathbf{s} \rightarrow \mathbf{s}') = \frac{1}{\pi}\int_0^{\pi/2} \frac{d\theta}{\det\left(\mathbf{I} + (8\sigma^2\sin^2\theta)^{-1}\mathbf{\Sigma}\mathcal{X}_{\mathbf{A}}^H(\mathbf{e})\mathcal{X}_{\mathbf{A}}(\mathbf{e})\right)}.$$

Now, our design problem can be stated as:

Problem 1: Find a matrix $\mathbf{A}$ that minimizes the worst-case average pair-wise error probability $\mathrm{P}_{\mathbf{A}}(\mathbf{s} \rightarrow \mathbf{s}')$, subject to the power constraint, $\mathrm{tr}\left(\mathbf{A}^H\mathbf{A}\right) \leq M$, i.e.,

$$\mathbf{A}_{\mathrm{opt}} = \arg \min_{\mathrm{tr}(\mathbf{A}^H\mathbf{A}) \leq p} \ \max_{\substack{\mathbf{s}, \mathbf{s}' \in \mathcal{S}^M \\ \mathbf{s}' \neq \mathbf{s}'}} P_{\mathbf{A}}(\mathbf{s} \rightarrow \mathbf{s}'),$$

where $\mathcal{S}^M = \mathcal{S} \times \mathcal{S} \cdots \times \mathcal{S}$. ■

To solve this problem, we introduce the following definition and lemmas.

Definition 2: Define the minimum distance of the constellation $\mathcal{S}$ as $d_{\min}(\mathcal{S}) = \min_{s \neq s', s, s' \in \mathcal{S}} |s - s'|$. ■

The following lemma relates a measure of the distance between matrices $\mathcal{X}_{\mathbf{I}_M}(\mathbf{s})$ and $\mathcal{X}_{\mathbf{I}_M}(\mathbf{s}')$ to $d_{\min}(\mathcal{S})$. For notational convenience, let $\mathcal{X}_{\mathbf{I}_M}(\mathbf{e}, \{i_1, i_2, \cdots, i_n\})$ denote the matrix that remains after the columns of $\mathcal{X}_{\mathbf{I}_M}(\mathbf{s})$ indexed by $\{i_1, i_2, \cdots, i_n\}$ have been removed.

Lemma 2: For any nonzero vector $\mathbf{e}$, we have

$$\det\left(\mathcal{X}_{\mathbf{I}_M}^H(\mathbf{e}, \{i_1, i_2, \cdots, i_n\})\mathcal{X}_{\mathbf{I}_M}(\mathbf{e}, \{i_1, i_2, \cdots, i_n\})\right) \geq d_{\min}^{2(M-n)}(\mathcal{S})$$

for $n = 0, 1, \cdots, M-1$, where the equality holds if and only if $\mathbf{s}$ and $\mathbf{s}'$ are neighbours; i.e., if and only if $\|\mathbf{e}\| = d_{\min}(\mathcal{S})$. ■

The following lemma provides a lower bound on the worst-case average pair-wise error probability.

Lemma 3: Let $\mathbf{D} = \mathrm{diag}(d_1, d_2, \cdots, d_M)$ with $d_n > 0$ for $n = 1, 2, \cdots, M$. Then, for any nonzero vector $\mathbf{e}$, the following inequality holds

$$\det\left(\mathbf{D} + \mathcal{X}_{\mathbf{I}_M}^H(\mathbf{e})\mathcal{X}_{\mathbf{I}_M}(\mathbf{e})\right) \geq \det\left(\mathbf{D} + d_{\min}^2(\mathcal{S})\mathbf{I}_M\right) \qquad (19)$$

with equality holding if and only if $\mathbf{s}$ and $\mathbf{s}'$ are neighbours. ■

The proofs of Lemmas 2 and 3 are omitted because of space limitation, which will be provided in somewhere else. We now state another main result.

Theorem 2: Let the eigenvalue decomposition of $\mathbf{\Sigma}$ be $\mathbf{\Sigma} = \mathbf{V}\mathbf{\Lambda}\mathbf{V}^H$, where $\mathbf{V}$ is an $M \times M$ unitary matrix and $\mathbf{\Lambda} = \mathrm{diag}(\lambda_1, \lambda_2, \cdots, \lambda_M)$ with $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_M > 0$. Then, we have the following statements.

1) An optimal solution for Problem 1 is given by $\mathbf{A}_{\mathrm{opt}} = \widetilde{\mathbf{\Gamma}}\mathbf{V}^H$, where the optimal $\widetilde{\mathbf{\Gamma}} = \mathrm{diag}(\widetilde{\mu}_1, \widetilde{\mu}_2, \cdots, \widetilde{\mu}_M)$ can be obtained by solving the following optimization problem:

$$\widetilde{\mathbf{\Gamma}} = \arg \min_{\mathrm{tr}(\mathbf{\Gamma}) \leq M} J_M\left(\mathbf{\Lambda\Gamma}, \frac{d_{\min}^2(\mathcal{S})}{8\sigma^2}\right), \qquad (20)$$

where $\mathbf{\Gamma} = \mathrm{diag}(\mu_1, \mu_2, \cdots, \mu_M)$ and $J_M(\mathbf{\Lambda\Gamma}, \varepsilon)$ denote the integral

$$J_M(\mathbf{\Lambda\Gamma}, \varepsilon) = \frac{1}{\pi}\int_0^{\pi/2}\prod_{k=1}^{M}\left(1 + \frac{\varepsilon\lambda_k\mu_k^2}{\sin^2\theta}\right)^{-1} d\theta \quad \text{for } \varepsilon > 0.$$

2) For such an optimal solution $\mathbf{A}_{\mathrm{opt}}$, the worst case pair-wise error probability achieves the lower bound; i.e.,

$$\max_{\mathbf{s}, \mathbf{s}' \in \mathcal{S}^N, \mathbf{s} \neq \mathbf{s}'} P_{\mathbf{A}_{\mathrm{opt}}}(\mathbf{s} \rightarrow \mathbf{s}') = J_M\left(\mathbf{\Lambda\Gamma}, \frac{d_{\min}^2(\mathcal{S})}{8\sigma^2}\right). \quad (21)$$

In addition, $P_{\mathbf{A}_{\mathrm{opt}}}(\mathbf{s} \rightarrow \mathbf{s}') = J_M\left(\mathbf{\Lambda\Gamma}, \frac{d_{\min}^2(\mathcal{S})}{8\sigma^2}\right)$ if and only if $\|\mathbf{s} - \mathbf{s}'\| = d_{\min}(\mathcal{S})$. ■

Proof: First we establish an lower bound on the worst case average pair-wise error probability. For an arbitrarily given positive integer $1 \leq m \leq N$, set $|e_m| = d_{\min}(\mathcal{S})$ and

$e_k=0, k=1,2,\cdots,N, k\neq m$. Then, exploiting the results in [35], [36], we get

$$\begin{aligned}
&\det\big(\mathbf{I}_M+\frac{1}{8\sigma^2\sin^2\theta}\mathbf{\Sigma}\mathcal{X}_{\mathbf{A}}^H(\mathbf{e})\mathcal{X}_{\mathbf{A}}(\mathbf{e})\big)\\
&=\det\big(\mathbf{I}_M+\frac{d_{\min}^2(\mathcal{S})}{8\sigma^2\sin^2\theta}\mathbf{\Sigma}\mathbf{A}^H\mathbf{A}\big)\\
&\leq\prod_{k=1}^{M}\big(1+\frac{d_{\min}^2(\mathcal{S})}{8\sigma^2\sin^2\theta}\mu_k^2\lambda_k\big) \qquad (22)
\end{aligned}$$

where equality in (22) holds if $\mathbf{A}=\mathbf{\Gamma}\mathbf{V}^H$. Therefore, the worst case average pair-wise error probability is lower bounded by

$$\max_{\mathbf{s},\mathbf{s}'\in\mathcal{S}^K,\mathbf{s}\neq\mathbf{s}'}P_{\mathbf{A}}(\mathbf{s}\to\mathbf{s}')\geq J_M\left(\mathbf{\Gamma}\mathbf{\Lambda},\,\frac{d_{\min}^2(\mathcal{S})}{8\sigma^2}\right).$$

In the following, we establish an upper bound of the worst case average pair-wise error probability. Notice that when $\mathbf{A}=\mathbf{\Gamma}\mathbf{V}^H$, we have

$$\begin{aligned}
&\det\left(\mathbf{I}_M+\frac{1}{8\sigma^2\sin^2\theta}\mathbf{\Sigma}\mathbf{X}_{\mathbf{A}}^H(\mathbf{e})\mathbf{X}_{\mathbf{A}}(\mathbf{e})\right)\\
&=\left(\frac{1}{8\sigma^2\sin^2\theta}\right)^M\det(\mathbf{\Lambda})\det\left(\mathbf{D}+\mathbf{X}_{\mathbf{I}_M}^H(\mathbf{e})\mathbf{X}_{\mathbf{I}_M}(\mathbf{e})\right), \qquad (23)
\end{aligned}$$

where $\mathbf{D}=(8\sigma^2\sin^2\theta)\mathbf{\Lambda}^{-1}\mathbf{\Gamma}^{-2}$. Using Lemma 3 we obtain that for any nonzero vector $\mathbf{e}$ and nonzero θ in the interval $[0,\pi/2]$,

$$\det\big(\mathbf{D}+\mathbf{X}_{\mathbf{I}_M}^H(\mathbf{e})\mathbf{X}_{\mathbf{I}_M}(\mathbf{e})\big)\geq\prod_{k=1}^{M}\big(\frac{8\sigma^2\sin^2\theta}{\mu_k^2\lambda_k}+d_{\min}^2(\mathcal{S})\big). \quad (24)$$

Here, the equality holds if and only if $\mathbf{s}$ and $\mathbf{s}'$ are neighbor points, i.e., $\|\mathbf{s}-\mathbf{s}'\|=d_{\min}(\mathcal{S})$. Therefore, combining (23) with (24) yields

$$\det\big(\mathbf{I}_M+\frac{1}{8\sigma^2\sin^2\theta}\mathbf{X}_{\mathbf{A}_{\rm opt}}^H(\mathbf{e})\mathbf{X}_{\mathbf{A}_{\rm opt}}(\mathbf{e})\mathbf{\Lambda}\big)\geq\prod_{k=1}^{M}\big(1+\frac{d_{\min}^2(\mathcal{S})\mu_k^2\lambda_k}{8\sigma^2\sin^2\theta}\big).$$

This results in

$$\max_{\mathbf{s},\mathbf{s}'\in\mathcal{S}^N,\,\mathbf{s}\neq\mathbf{s}'}P_{\mathbf{A}_{\rm opt}}(\mathbf{s}\to\mathbf{s}')\leq J_M\left(\mathbf{\Lambda}\mathbf{\Gamma},\,\frac{d_{\min}^2(\mathcal{S})}{8\sigma^2}\right), \qquad (25)$$

where the equality in (25) holds if and only if $\|\mathbf{s}-\mathbf{s}'\|=d_{\min}(\mathcal{S})$. Combining (25) with Lemma 2 yields

$$\max_{\substack{\mathbf{s},\mathbf{s}'\in\mathcal{S}^N\\ \mathbf{s}\neq\mathbf{s}'}}P_{\mathbf{A}_{\rm opt}}(\mathbf{s}\to\mathbf{s}')=J_M\left(\mathbf{\Lambda}\mathbf{\Gamma},\,\frac{d_{\min}^2(\mathcal{S})}{8\sigma^2}\right),$$

and hence. Statements 1 and 2 of the theorem are true. □

Corollary 2: If the pair-wise error probability is upper bounded by Chernoff bound, then, the optimization problem in (20) can be relaxed by the following optimization problem

$$\widetilde{\mathbf{A}}_{\rm opt}=\arg\max_{{\rm tr}(\widetilde{\mathbf{A}}^H\widetilde{\mathbf{A}})\leq M}\det\left(\mathbf{I}+\frac{d_{\min}^2(\mathcal{S})}{8\sigma^2}\widetilde{\mathbf{A}}\mathbf{\Sigma}\widetilde{\mathbf{A}}^H\right). \quad (26)$$

Therefore, its optimal solution $\widetilde{\mathbf{A}}_{\rm opt}=\widetilde{\mathbf{\Gamma}}\mathbf{V}^H$, where $\widetilde{\mathbf{\Gamma}}={\rm diag}(\widetilde{\mu}_1,\widetilde{\mu}_2,\cdots,\widetilde{\mu}_M)$ can be obtained by employing the water-filling strategy [36],

$$\widetilde{\mu}_m=\sqrt{\left[\frac{1}{\overline{M}}\left(M+\frac{8\sigma^2}{d_{\min}^2(\mathcal{S})}\sum_{\ell=1}^{\overline{M}}\frac{1}{\lambda_\ell}\right)-\frac{1}{\lambda_n}\right]_+}, \qquad (27)$$

where notation $[x]_+$ denotes $\max(x,0)$. Particularly when $\mathbf{\Sigma}=\mathbf{I}_M$, any $M\times M$ unitary matrix is one of the optimal solution of Problem 1. ■

V. Optimal diversity-vs-multiplexing tradeoff

In this section we assume that the channel coefficients are independent; i.e., $\mathbf{\Sigma}=\mathbf{I}$, and show that the linear ZF receiver enables to achieve diversity-vs-multiplexing tradeoff [1] for our Toeplitz code. In the following, the notion is adopted from [1]. Consider the square QAM constellation with size $\sqrt{D}$ per real dimension. Then, the transmission data rate R is $R=\frac{N}{T}\log_2 D$, where $\frac{N}{T}$ is the symbol rate with N being the length of the transmitted signal vector $\mathbf{s}$. Define the multiplexing gain r as [1] $R=r\log_2{\rm SNR}$. Notice that we always have $0\leq r\leq 1$ since the system has only one receiver antenna. Hence, we obtain $D={\rm SNR}^{Tr/N}$. Now, the averaged transmission energy E_s for square QAM signal is [37] $E_s=\frac{2}{3}(D-1)$. Hence the averaged transmission energy per block is $E_{s{\rm bl}}=\frac{2}{3}(D-1)MN$. Given σ^2 being the noise variance at the receiver antenna, the averaged noise power per block is $\sigma_{\rm bl}^2=\sigma^2T$, where $T=M+N-1$, the block length in time dimension. Therefore, we have

$$\sigma^2=\frac{2(D-1)MN}{3T{\rm SNR}}\doteq\frac{2DMN}{3T{\rm SNR}}=\frac{2MN}{3T}{\rm SNR}^{\frac{Tr}{N}-1} \quad (28)$$

where the second step comes from the assumption of high SNR. Now, consider the SEP for square QAM signals with ZF receiver. Using the upper bound (17), we have

$${\rm E}\left[{\rm P}(\mathbf{h})\right]\leq\frac{D-1}{D}\det((1+a_{\rm QAM}C)\mathbf{I})^{-1}\leq\ C^{-M}a_{\rm QAM}^{-M}.$$

The second inequality comes from the assumption of high SNR. On the other hand, notice that

$$a_{\rm QAM}=\frac{3}{2}\frac{E_s}{\sigma^2(D-1)}=\sigma^{-2}=\frac{3T}{2MN}\ {\rm SNR}^{1-Tr/N}.$$

Therefore, we obtain ${\rm E}\left[{\rm P}(\mathbf{h})\right]\leq\left(\frac{3TC}{2MN}\right)^{-M}{\rm SNR}^{\frac{MT}{N}r-M}$. Hence, the diversity order is $d_{\rm Toe}(r)=M-\frac{MT}{N}r$. Define $\delta=\frac{M-1}{N}$ which is nonnegative, then we have $d_{\rm Toe}(r)=M(1-r)-\delta Mr=d_{\rm opt}(r)-\delta Mr$, where $d_{\rm opt}(r)=M(1-r)$ is the optimal diversity-vs-multiplexing tradeoff proposed by [1] for a MISO system. For any small δ, we can always choose $N=\lceil\frac{M-1}{\delta}\rceil+1$, where $\lceil\cdot\rceil$ is the integer part of a quantity. Hence, $d_{\rm Toe}(r)$ is the δ approximation of $d_{\rm opt}(r)$. Therefore, we can say that the ZF receiver is able to approach the optimal diversity-vs-multiplexing tradeoff for the proposed Toeplitz code.

VI. Simulations

To demonstrate our Toeplitz STBC code, we consider a coherent MISO system with three transmitter antennas and a single receiver antenna. Fig. 1 shows the average bit error performance comparison of our code with the orthogonal STBC with a half symbol rate [9], in which $T=8$ and the transmitted signal is 16-QAM for our code and 64-QAM for the orthogonal STBC.

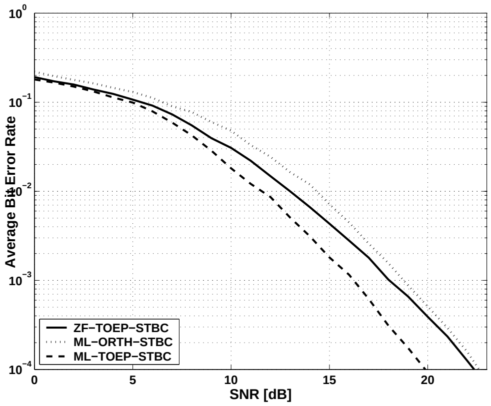

Fig. 1. The error performance comparison of our code with the orthogonal STBC. The dash curve and solid curve denote the ML receiver and ZF receiver for our code, respectively. The dotted curve denotes the ML detector for the orthogonal STBC.

VII. Conclusion

In this paper, we designed Toeplitz linear space time block codes that minimize the exact worst case average pair-wise error probability for coherent correlated MISO flat fading channels. For the independent MISO flat fading channels, we proved that our proposed codes can approach the optimal diversity-vs-multiplexing tradeoff with a linear ZF receiver when the number of channel uses is large.

References

[1] L. Zheng and D. N. C. Tse, "Diversity and multiplexing: A fundamental tradeoff in multiple-antennas channels," *IEEE Trans. Inform. Theory*, vol. 49, pp. 1073–1096, May 2004.

[2] A. Narula, M. D. Trott, and G. W. Wornell, "Performance limits of coded diversity methods for transmitter antenna arrays," *IEEE Trans. Inform. Theory*, vol. 45, pp. 2418–2433, Nov. 1999.

[3] A. L. Moustakas, "Communication through a diffusive medium: Coherence and capacity," *Science*, vol. 287, pp. 287–290, Jan. 2000.

[4] E. Visotsky and U. Madhow, "Space-time transmit precoding with imperfect feedback," *IEEE Trans. Inform. Theory*, vol. 47, pp. 2632–2639, Sept. 2001.

[5] S. Vishwanath, S. A. Jafar, and A. Goldsmith, "Channel capacity and beamforing for multiple transmit and receive antennas with covariance feedback," in *ICC*, (Helsinki Finland), 2001.

[6] S. A. Jafar and A. Goldsmith, "On opimality of beaforming for multiple antenna systems with imperfect feedback," in *Proceedings of IEEE International Symposium on Inform. Theory*, (Washington DC), pp. 321–321, June 2001.

[7] G. B. Giannakis and S. Zhou, "Optimal tranmitter diversity precoders for random fading channel," in *Proc. of Globecom Conf.*, (San Francisco, CA), pp. 1839–1843, Nov. 2000.

[8] S. M. Alamouti, "A simple transmitter diversity scheme for wireless commincations," *IEEE J. Select. Areas Commun*, vol. 16, pp. 1451–1458, Oct. 1998.

[9] V. Tarokh, H. Jafarkhani, and A. R. Calderbank, "Space-time block codes from orthogonal designs," *IEEE Trans. Inform. Theory*, vol. 45, pp. 1456–1467, July. 1999.

[10] O. Tirkkonen and A. Hottinen, "Square-matrix embeddable space-time codes for complex signal constellations," *IEEE Trans. Inform. Theory*, vol. 48, pp. 1122–1126, Feb. 2002.

[11] X.-B. Liang, "Orthogonal designs with maximal rates," *IEEE Trans. Inform. Theory*, vol. 49, pp. 2468–2503, Oct. 2003.

[12] X.-B. Liang and X.-G. Xia, "Nonexistence of rate one space-time blocks from generalized complex linear processing orthogonal designs for more than two transmit antennas," in *International Symposium on Inform. Theory*, (Washington DC), June. 2001.

[13] W. Su and X.-G. Xia, "Two generalized complex orthogonal space-time block codes of rares 7/11 and 3/5 for 5 and 6 transmit antennas," in *International Symposium on Inform. Theory*, (Washington DC), June. 2001.

[14] H. Wang and X.-G. Xia, "Upper bounds of rates of complex orthogonal space-time block codes," in *International Symposium on Inform. Theory*, (Laussane, Switzerland), June. 2002.

[15] S. Sandhu and A. J. Paulraj, "Space-time block coding: A capacity perspective," *IEEE Commun. Letters*, vol. 4, pp. 384–386, Dec. 2000.

[16] B. Hassibi and B. M. Hochwald, "High-rate codes that are linear in space and time," *IEEE Trans. Inform. Theory*, vol. 50, pp. 1804–1824, Jul. 2002.

[17] H. E. Gamal and M. O. Damen, "Universal space-time coding," *IEEE Trans. Inform. Theory*, vol. 49, pp. 1097–1119, May. 2003.

[18] X.-Ma and G. B. Giannakis, "Full-diversity full rate complex-field space-time coding," *IEEE Trans. Signal Processing*, vol. 51, pp. 2917–2930, Nov. 2003.

[19] B. A. Sethuraman, B. S. Rajan, and V. Shashidhar, "Full-diversity, high rate space-time block codes from division algebras," *IEEE Trans. Inform. Theory*, vol. 49, pp. 2596–2616, Oct. 2003.

[20] J.-K. Zhang, J. Liu, and K. M. Wong, "Trace-orthonormal full diversity cyclotomic linear dispersion codes," in *Proceedings IEEE International Symposium on Information Theory*, (Chicago), June 2004.

[21] H. Yao and G. W. Wornell, "Achieving the full MIMO diversity-vs-multiplexing frontier with rotation-based space-time codes," in *41th Annual Allerton Conf. on Comm. Control, and Comput.*, (Monticello, IL), Oct. 2003.

[22] P. Dayal and M. K. Varanasi, "An optimal two tansmite antenna space-time code and its stacked extensions," in *Asilomar Conf. on Signals, Systems and Computers*, (Monterey, CA), Nov. 2003.

[23] J. C. Belfiore and G. Rekaya, "Quaternionic lattices for space-time coding," in *Proceedings IEEE of ITW2003*, (Paris, France), March 2003.

[24] J. C. Belfiore, G. R. Rekaya, and E. Viterbo, "The Golden code: A 2×2 full rate space-time code with non-vanishing determinants," in *Proceedings IEEE International Symposium on Information Theory*, (Chicago), June 2004.

[25] G. Rekaya, J. C. Belfiore, and E. Viterbo, "Algebraic 3×3, 4×4 and 6×6 space-time codes with non-vanishing determiants," in *ISITA*, (Parma, Italy), Oct. 2004.

[26] G. Wang and X.-G. Xia, "Optimal multi-layer cyclotomic space-time code designs," in *Proceedings IEEE International Symposium on Information Theory*, (Chicago), June 2004.

[27] J.-K. Zhang, G.-Y. Wang, and K. M. Wong, "Optimal norm form integer space-time codes for two antenna MIMO systems," in *Int. Conf. Acoust., Speech, Signal Process.*, (Philadelphia, USA), March 2005.

[28] H. E. Gamal, G. Caire, and M. O. Damen, "Lattice coding and decoding achieve the optimal diversity–multiplexing tradeoff of MIMO channels," *IEEE Trans. Inform. Theory*, vol. 50, pp. 968–985, June 2004.

[29] U. Erez and R. Zamir, "Lattice decoding can achieve $\frac{1}{2}\log(1+\mathrm{snr})$ on the AWGN channel using nested codes," *IEEE Trans. Inform. Theory*, submitted for publication.

[30] H.-F. Lu and P. V. Kumar, "Rate-diversity tradeoff of space-time codes with fixed alphabet and optimal constructions for PSK modulation," *IEEE Trans. Inform. Theory*, vol. 49, pp. 2747–2751, Oct. 2003.

[31] G. D. Forney Jr., "Maximum-likelihood sequence estimation of digital sequences in the presence of intersymbol interference," *IEEE Trans. Inform. Theory*, vol. 18, pp. 363–378, March 1972.

[32] R. A. Horn and C. R. Johnson, *Matrix Analysis*. Cambridge, MA: Cambridge University Press, 1985.

[33] M. K. Simon and M.-S. Alouini, "A unified approach to the performance analysis of digital communication over generalized fading channels," *Proc. of the IEEE*, vol. 86, pp. 1860–1877, Sept. 1998.

[34] G. D. Forney and G. U. Ungerboeck, "Modulation and coding for linear Gaussian channel," *IEEE Trans. Inform. Theory*, vol. 44, pp. 2384–2415, May 1998.

[35] H. S. Witsenhausen, "A determinant maximization problem occurring in the theory of data," *SIAM J. App. Math.*, vol. 29, pp. 515–522, Nov 1975.

[36] T. M. Cover and J. A. Thomas, *Elements of Information Theory*. New York: John Wiley and Sons, 1991.

[37] J. G. Proakis, *Digital Comunications*. Bejing: Prentice-Hall, Inc, 1996.