

SPACE-TIME LINEAR DISPERSION CODES BASED ON OPTIMAL ALGORITHMS

Yan Liang, Wen Chen, Yan She

Electronic Information and Electrical Engineering
Electronic Engineering
Jiaotong University, Minhang, Shanghai, China
{louiseliang; wenchen}@sjtu.edu.cn}

ABSTRACT

Multiple-antenna systems that operate at high rates require simple yet effective space-time transmission schemes. Linear dispersion (LD) codes transmits substreams of data in linear combinations over space and time. The traditionally LD codes are mainly designed to optimize the mutual information which cannot guarantee good error performance. In contrast, using the simulation-based optimization method can minimize the block error rate (BLER), but failed to optimize the channel capacity to the fullest extent. In this paper, a new design algorithm will be presented which would achieve both lower BLER and higher capacity by adding some extra qualifications in the simulation process. Simulation results show that the proposed method has a better performance over a large range of SNRs than the earlier proposed methods.

Index Terms— MIMO communication systems, mutual information, linear dispersion, space-time block coding, simulation based optimization.

1. INTRODUCTION

Wireless communication systems with multiple antennas at both the transmitter and receiver have the potential to reliable transmission at high data rates and have attracted extensive interest and research. Before linear dispersion codes (LDC) was presented by Hassibi and Hochwald [1], people focused mainly on Bell Labs Layered Space-Time architecture (BLAST) [2] or schemes like orthogonal space-time block codes (STBCs) [3] [4]. The former one can reach the rates of tens of bits per second per hertz but suffers from its inability in working with fewer receive antennas than transmit antennas due to its decoding scheme [5]. On the contrary, the latter approach can have a much better fading resistance and simpler decoding but perform poorly at high data rates or with many antennas. Linear dispersion codes play a role of combining the advantage of both BLAST and STBCs which pay more attention on the capacity of the channel and mutual information.

Supported by NSF China #60672067, NSF Shanghai #062R14041 and SJTU #T225403001

In [6], Gamal proposed a scheme called threaded algebraic space-time (TAST) coding. This scheme can achieve full diversity and full rate using the basic frame of LD codes [1]. Comparing with the traditional LD codes which focus mainly on the optimization of mutual information, its design based on the worst case of pairwise error probability (PEP). Also, PEP expressions under quasi-static fading, fast fading, block fading, as well as arbitrarily temporally correlated fading, under Rayleigh and Rician conditions were discussed in [7]. LD codes are typically designed for flat fading channels, but in [8] Bölskei extends LDC designs for frequency selective fading channels using Orthogonal Frequency Division Multiplexing (OFDM). On the other hand, in [9]-[11] Wang proposed a series optimization design based on the simulation and gradient estimation. However, focusing only on the minimum block error rate (BLER) in the LD design can not guarantee a reasonable tradeoff between different perspectives, viz., MMSE, mutual information, and average pairwise error probability (PEP) [12]. This paper will provide a new method that provides a good tradeoff between the above mentioned perspectives.

2. MIMO SYSTEM WITH LINEAR DISPERSION CODES

We consider a MIMO system with M transmit antennas, N receive antennas, and an interval of T symbols available to us during which the propagation channel is constant and known to the receiver.

From [7], we know the input-output relation is given by

$$Y = \sqrt{\frac{\rho}{M}} SH + W, \quad (1)$$

where Y is the $T \times N$ matrix of the received signal, S is the $T \times M$ matrix of the transmitted signal, W is the $T \times N$ matrix of the additive white Gaussian noise, and H is the $M \times N$ MIMO channel matrix. We assume that the data sequence has been broken into Q substreams and s_1, \dots, s_Q are the complex symbols chosen from an arbitrary, say r-PSK or r-QAM, constellation. The code is completely determined by the set

of $T \times M$ dispersion matrices $\{\mathcal{A}_q, \mathcal{B}_q\}$. By [1], write the transmitted signal \mathbf{S} in the form $S = \sum_{q=1}^Q (\alpha_q \mathcal{A}_q + j\beta_q \mathcal{B}_q)$, where the real scalars $\{\alpha_q, \beta_q\}$ are determined by

$$s_q = \alpha_q + j\beta_q, \quad \text{for } q = 1, \dots, Q. \quad (2)$$

Without loss of generality, we assume that $\alpha_1, \dots, \alpha_Q$ and β_1, \dots, β_Q have the variance 0.5 and are uncorrelated. Thus, s_1, \dots, s_Q are unit-variance and uncorrelated. So the transmit signal \mathbf{S} can be normalized such that $E\text{tr}(SS^*) = TM$, where the $\text{tr}(SS^*)$ denotes the trace of SS^* and $E(\cdot)$ is the expectation. This induces the following normalization on matrices $\{\mathcal{A}_q, \mathcal{B}_q\}$:

$$\sum_{q=1}^Q [\text{tr}(\mathcal{A}_q^* \mathcal{A}_q) + \text{tr}(\mathcal{B}_q^* \mathcal{B}_q)] = 2TM, \quad (3)$$

where $(\cdot)^*$ denotes the Hermitian transpose. If r denotes the constellation size of symbol s_q , then the transmission rate of the code is $R = \frac{Q}{T} \log_2 r$ bits/channel. We then gather the equations in real and imaginary part to form the single real system of equations:

$$\begin{bmatrix} y_{R,1} \\ y_{I,1} \\ \vdots \\ y_{p,N} \\ y_{I,N} \end{bmatrix} = \sqrt{\frac{\rho}{M}} \mathcal{H} \begin{bmatrix} \alpha_1 \\ \beta_1 \\ \vdots \\ \alpha_Q \\ \beta_Q \end{bmatrix} + \begin{bmatrix} w_{R,1} \\ w_{I,1} \\ \vdots \\ w_{R,N} \\ w_{I,N} \end{bmatrix}, \quad (4)$$

where $y_{R,i}$ and $y_{I,i}$ are, respectively, the real and imaginary parts of y_i , so are α_i, β_i of s_i and $w_{R,i}, w_{I,i}$ of w_i .

Suppose that

$$A_q = \begin{bmatrix} \mathcal{A}_{R,q} & -\mathcal{A}_{I,q} \\ \mathcal{A}_{I,q} & \mathcal{A}_{R,q} \end{bmatrix}, \quad B_q = \begin{bmatrix} -\mathcal{B}_{I,q} & -\mathcal{B}_{R,q} \\ \mathcal{B}_{R,q} & -\mathcal{B}_{I,q} \end{bmatrix}, \quad (5)$$

where $\mathcal{A}_q = \mathcal{A}_{R,q} + j\mathcal{A}_{I,q}$ and $\mathcal{B}_q = \mathcal{B}_{R,q} + j\mathcal{B}_{I,q}$.

At this time, the equivalent $2NT \times 2Q$ real channel matrix is given by:

$$\mathcal{H} = \begin{bmatrix} A_1 h_1 & B_1 h_1 & \cdots & A_Q h_1 & B_Q h_1 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ A_1 h_N & B_1 h_N & \cdots & A_Q h_N & B_Q h_N \end{bmatrix},$$

where $h_i = [H_{R,i}^*, H_{I,i}^*]^*$. H_i is the i th column vector of the physical channel matrix H . $H_{R,i}$ and $H_{I,i}$ are respectively the real and imaginary part of H_i .

3. OPTIMAL DESIGN

3.1. Mutual Information

The traditional method to design LD codes is to analyze the channel capacity $C(\rho, M, N)$ assuming the capacity achieving standard circular Gaussian distribution

$$C(\rho, M, N) = E \log \det(I_{2NT} + \frac{\rho}{M} \mathcal{H} \mathcal{H}^*), \quad (6)$$

and find out proper $\{A_q, B_q\}$ that can optimize the $C(\rho, M, N)$, subject to

$$\text{tr}(A_q^* A_q), \text{tr}(B_q^* B_q) \leq \frac{TM}{Q}, \quad \text{for } q = 1, \dots, Q, \quad (7)$$

By Jensen's inequality¹, the capacity can be expressed as

$$C(\rho, M, N) \leq \log \det\{I_{2NT} + \frac{\rho}{M} E(\mathcal{H} \mathcal{H}^*)\}. \quad (8)$$

Observe that

$$\mathcal{H} \mathcal{H}^* = \sum_{q=1}^Q [(I_N \otimes A_q) \underline{h} \underline{h}^* (I_N \otimes A_q^*) + (I_N \otimes B_q) \underline{h} \underline{h}^* (I_N \otimes B_q^*)],$$

where $\underline{h} = [h_1^*, \dots, h_N^*]^*$, and assume the coefficients of H are independent Gaussian random variables of unit variance. Since $E(\underline{h} \underline{h}^*) = I_{2M}$, we can get

$$E(\mathcal{H} \mathcal{H}^*) = \frac{1}{2} [I_N \otimes \sum_{q=1}^Q (A_q A_q^* + B_q B_q^*)], \quad (9)$$

and also

$$C(\rho, M, N) \leq \log \det\{I_{2NT} + \frac{\rho}{2M} [I_N \otimes \sum_{q=1}^Q (A_q A_q^* + B_q B_q^*)]\}.$$

From Hadamard's inequality, it is known that for any positive definite matrix X : $\det(X) \leq \prod_{ii} x_{ii}$, where the x_{ii} are the diagonal entries of X . Moreover, the equation holds if and only if X is diagonal. Let

$$X = I_{2NT} + \frac{\rho}{2M} [I_N \otimes \sum_{q=1}^Q (A_q A_q^* + B_q B_q^*)]. \quad (10)$$

Using Hadamard's inequality to maximize the upper bound of the Capacity equation, the optimal choice of the set $\{A_q, B_q\}$ follows from (3) and (10)

$$\sum_{q=1}^Q (A_q A_q^* + B_q B_q^*) = 2M I_{2T}. \quad (11)$$

Thus, we can write the maximized bound on the mutual information as

$$I(s; y) \leq N \log(1 + \rho). \quad (12)$$

The conclusion which we get in (11) can maximize the upper bound of capacity and will be used as a qualification in the process of simulation in subsection C to achieve a higher capacity.

¹If f is a strictly convex function and X is a random variable, the $E\{f(X)\} \leq f\{E(X)\}$ with equality if $X = E(X)$ with probability 1.

3.2. Pairwise Error Probability

In LDC design, minimizing average pairwise error probability (PEP) is shown to be numerically difficult for high rate systems[1]. Rather, LDC design was achieved by formulating a power-constrained optimization problem based on mutual information.

Let $P(s \rightarrow s'|\mathcal{H})$ denotes the probability that, conditioned on a specific equivalent channel realization \mathcal{H} , the receiver selects the codeword s' instead of the sent codeword s , then the average PEP can be defined as[1]

$$E_{S,S'}\{P(s \rightarrow s')|\mathcal{H}\} = \frac{1}{4\pi} \int_{-\infty}^{\infty} \frac{d\omega}{\omega^2 + 0.25} \prod_{q=1}^{2Q} \frac{1}{\sqrt{1 + \frac{2\rho\sigma_q(\omega^2+0.25)}{M}}},$$

where σ_q denotes the eigenvalues of $\mathcal{H}\mathcal{H}^*$. If the SNR is sufficiently high, the above equation can be simplified to

$$E_{S,S'}\{P(s \rightarrow s')|\mathcal{H}\} \leq \eta_r \left(\frac{\rho}{M}\right)^{-\frac{r}{2}} \prod_{q=1}^r \sigma_q^{-\frac{1}{2}}, \quad (13)$$

where $r = \text{Rank}(\mathcal{H})$.

As r is the rank of matrix \mathcal{H} , its value can be determined by the design of A_q, B_q . Since $\prod_{q=1}^r \sigma_q = \det(\mathcal{H})$, we can have

$$\prod_{q=1}^r \sigma_q^{-\frac{1}{2}} = \begin{cases} \det(\mathcal{H}\mathcal{H}^*)^{-\frac{1}{2}}, & \text{if } Q \geq NT, \\ \det(\mathcal{H}^*\mathcal{H})^{-\frac{1}{2}}, & \text{if } Q \leq NT. \end{cases} \quad (14)$$

To minimize the PEP, we can assume $r = \min(Q, NT)$. So the inequality can be simplified as

$$PEP_{av} = \begin{cases} \eta_{2NT} \left(\frac{\rho}{M}\right)^{-NT} E\{\det(\mathcal{H}\mathcal{H}^*)^{-\frac{1}{2}}\} & \text{if } Q \geq NT, \\ \eta_{2Q} \left(\frac{\rho}{M}\right)^{-NT} E\{\det(\mathcal{H}^*\mathcal{H})^{-\frac{1}{2}}\} & \text{if } Q \leq NT. \end{cases}$$

Then the goal to minimize the PEP turns out to be minimizing the value of $\det(\mathcal{H}\mathcal{H}^*)^{-\frac{1}{2}}$. We can see from the above section that the optimal set $\{A_q, B_q\}$: $\sum_{q=1}^Q (A_q A_q^* + B_q B_q^*) = 2M I_{2T}$ can just satisfied the needs here in minimizing the PEP_{av} .

3.3. Average Block Error Probability

Denote $H = [h_{1,1}^*, \dots, h_{N,1}^*, \dots, h_{1,M}^*, \dots, h_{N,M}^*]^*$ and denote the set of dispersion matrices as $\theta = \{A_{q,k}, B_{q,k}\}_{q=1}^Q, k = 1, 2, \dots, M\}$. Let $\gamma(y, s, h, \theta)$ denote the empirical BLEP for a given set of dispersion matrices θ . In [9] the authors proposed the average BLEP is obtained by

$$R(\theta) = E_{y,s,h|\theta}\{r(y, s, h, \theta)\}. \quad (15)$$

Take gradient, we have

$$\nabla_\theta R(\theta) = E_s E_H E_{y|s,H,\theta}\{r(y, s, H, \theta) \nabla_\theta \log p(y|x, h, \theta)\}.$$

The design procedure is as follows:

1. Generate samples Draw M symbol vectors s_1, s_2, \dots, s_M uniformly from the possible transmitted symbol set and then simulate M observations y_1, \dots, y_M . Using the decoding algorithm to compute the BLEP $\gamma(y, s, h, \theta)$.

2. Compute gradient estimation

$$g(\theta_n) = \frac{1}{M} \sum_{i=1}^M r(y_i, s_i, h_i, \theta_n) [\nabla_\theta \log p(y_i|s_i, h_i, \theta)]$$

3. Update dispersion matrices

$$\theta_{n+1} = \begin{cases} \theta_n - a_n g(\theta_n), & \theta \in \Theta, \\ \{d_k A_{q,k}, d_k B_{q,k}\}_{q=1}^Q, K = 1, \dots, M & \theta \notin \Theta, \end{cases}$$

where

$$d_k = \begin{cases} 1, & \text{if } L \leq 2TM_T, \\ \sqrt{\frac{2TM_T}{\sum_{q=1}^Q \text{Tr}(A_{q,k}^* A_{q,k} + B_{q,k}^* B_{q,k})}}, & \text{if } L \geq 2TM_T, \end{cases}$$

$$\text{and } L = \sum_{q=1}^Q \text{Tr}(A_{q,k}^* A_{q,k} + B_{q,k}^* B_{q,k}).$$

The set Θ is much stricter than it proposed in [9] due to the consideration of mutual information and PEP.

$$\Theta = \{\theta : \sum_{q=1}^Q (A_q A_q^* + B_q B_q^*) = 2M I_{2T}\} \quad (16)$$

4. SIMULATION

In this section, we give an examples to show the performance of the new codes searched by the simulation-based optimization method under the two extra qualifications. In the procedure, we randomly generate the dispersion matrices as the initial point to be used in the first iteration. From the experience, we can find that the codes generated with different initial conditions usually end up with very similar error performance.

In the following examples, we search the codes by selecting the SNR such that the bit error probability is around 10^{-2} .

Example: $M_T = 2, M_R = 2, K = 2, T = 3$, and $Q=2$ with MMSE decoding. We consider a scenario with two users each with two transmit antennas and two receiver antennas. Each user transmits $Q=2$ QPSK constellations over $T=3$ symbol intervals, resulting the overall spectral efficiency of $\frac{8}{3} \text{ bits/s/Hz}$. The LD code designed using the earlier simulation-based method is denoted as LD1 and the code designed by the stricter qualifications is denoted as Purposed Method. From Fig.1, we can see that the proposed method code performs similar to the LD1 code in the bit error probability under the MMSE decoding. But as we can see in Fig.2, the capacity of the proposed method code has been optimized compared to the LD1 code to improve that the two extra qualifications guarantee the capacity in the performance.

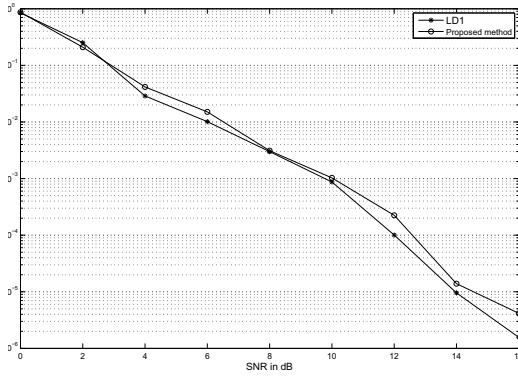


Fig. 1. Example: $M_T = 2, M_R = 2, K = 2, T = 3$, and $Q = 2$ with MMSE decoding

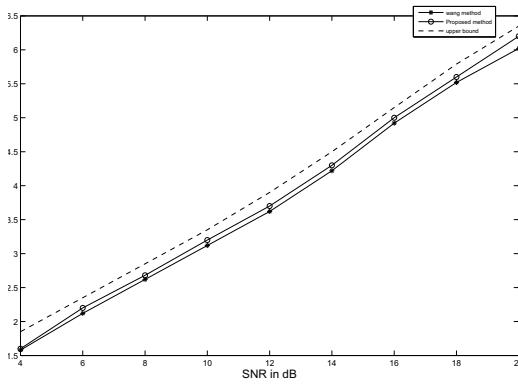


Fig. 2. Example: Actual channel capacity: the upper bound, the LD1 code and the proposed method code

5. CONCLUSION

In this paper, we have proposed to design LD codes by minimizing the union bound based on the exact PEP. Two extra qualifications for the simulation-based optimization approach are developed to optimize the LD codes by both minimizing the BLER and improving the channel capacity. The proposed two qualifications can also be applied as a universal method for other LD codes designs over wireless MIMO channels. Simulation results show that codes generated by the new algorithm generally outperform the codes designed only based on the mutual information or the average block error rate. Also, LD codes constructed from unitary coding matrices are asymptotically optimum from different design perspectives, viz., minimum mean square error (MMSE), mutual information and average pairwise error probability (PEP). Further improvement in designing the LD codes can be aimed at the symbol constellation, the detection algorithm or the spatial channel correlation structure.

6. REFERENCES

- [1] B. Hassibi, B. Hochwald, "High-rate codes that are linear in space and time," *IEEE Trans. Infrom. Theory*, vol. 48, no. 7, pp. 1804-1824, Jul. 2002.
- [2] G. D. Golden, G. J. Foschini R. A. Valenzuela and P. W. Wolniansky, "Detection algorithm and initial laboratory results using V-BLAST space-time communication architecture," *Electron. Lett.*, vol. 35, pp. 14-16, Jan. 1999.
- [3] S. M. Alamouti, "A Simple Transmit Diversity Technique for Wireless Communications," *IEEE Journ. Seclct Areas in Communications*, vol. 16, no. 8, pp. 1451-1457, Oct. 1998.
- [4] 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.
- [5] G. J. Foschini, G. D. Golden, R. A. Valenzuela, and P. W. Wolniansky, "Simplified processing for high spectral efficiency wireless communication employing multi-element arrays," *J.Select. Areas Commun.*, vol. 17, pp. 1841-1852, Nov. 1999.
- [6] H. E. Gamal and M. O. Damen, "Universal space-time coding," *IEEE Trans. Inf. Theory*, vol. 48, no. 5, pp. 1097-1119, May. 2003.
- [7] M. Uysal and C. Georghiades, "On the error performance analysis of space-time trellis codes," *IEEE Trans. Wireless Commun.*, vol. 3, no. 4, pp. 1118-1123, Jul. 2004.
- [8] H. Bölcseki, M. Borgmann, and A. J. Paulraj, "Impact of the propagation environment on the performance of space-frequency coded MIMO-OFDM," *IEEE J. Sel. Areas Commun.*, vol. 21, no. 3, pp. 427-439, Apr. 2003.
- [9] J. Wang, and X. Wang "Optimal Linear Space-Time Spreading for Multiuser MIMO Communications," *IEEE Journ. Secl. Areas in Commun.*, vol. 24, no. 1, pp. 113-120, Jan. 2006.
- [10] X. Wang, V. Krishnamurthy, and J. Wang "Stochastic Gradient Algorithms for Design of Minimum Error-Rate Linear Dispersion Codes in MIMO Wireless Systems," *IEEE Trans. on signal Processing*, vol. 54, no. 4, pp. 1242-1255, Apr. 2006.
- [11] J. Wang, X. Wang, and M. Madhian "On the Optimum Design of Space-Time Linear-Dispersion Codes," *IEEE Trans. on wireless Commun.*, vol. 4, no. 6, pp. 2928-2938, Nov., 2005.
- [12] A. Dogandžić, "Chernoff bounds on pairwise error probabilities of space-time codes," *IEEE Trans. Inf. Theory*, vol. 49, no. 5, pp. 1327-1336, May. 2003.