HIGH-RATE LINEAR SPACE-TIME CODES

B. Hassibi and B. Hochwald

Bell Labs, Lucent Technologies, Murray Hill, NJ

ABSTRACT

Multiple-antenna systems that operate at high rates require simple yet effective space-time transmission schemes to handle the large traffic volume in real time. V-BLAST, where every antenna transmits its own independent substream of data, has been shown to have good performance and simple encoding and decoding. Yet its drawbacks include its inability to work with fewer receive antennas than transmit antennas, and its absence of built-in spatial coding. On the other hand, there are many previously-proposed space-time codes that have good fading resistance and simple decoding, but generally poor performance at high data rates or with many antennas.

We propose a high-rate coding scheme that can handle any configuration of transmit and receive antennas and that subsumes both V-BLAST and many proposed space-time codes as special cases. The scheme transmits substreams of data in linear combinations over space and time and the codes are designed to optimize the mutual information between the transmitted and received signals. Because of their linear structure, the codes retain the decoding simplicity of V-BLAST, and because of their information-theoretic optimality, they possess many coding advantages.

1. INTRODUCTION AND MODEL

It is widely acknowledged that reliable fixed and mobile wireless transmission of video, data, and speech at high rates will be an important part of future telecommunications systems. One way to get high rates on a scattering-rich wireless channel is to use multiple transmit and/or receive antennas. To achieve the high data rates promised in theory [1], however, new approaches for space-time transmission are needed. One such approach is V-BLAST (Vertical Bell Labs Layered Space-Time) [2] which breaks the original data stream into substreams that are transmitted on the individual antennas. The receiver decodes the substreams using a sequence of nulling and cancelling steps.

Since then there has been considerable work on a variety of space-time transmission schemes and performance measures [3, 4, 5]. At very high rates and with a large number of antennas, many of these space-time codes suffer from complexity or performance difficulties. Although V-BLAST can

handle high data rates with reasonable complexity, the decoding scheme presented in [2] does not work with fewer receive than transmit antennas.

We present a space-time transmission scheme that has many of the coding and diversity advantages of previously designed codes, but also has the decoding simplicity of V-BLAST at high data rates. The codes work with arbitrary numbers of transmit and receive antennas. The codes break the data stream into substreams that are dispersed in linear combinations over space and time. We refer to them simply as LD codes (linear dispersion codes). They:

1. Subsume, as special cases, both V-BLAST [2] and the block codes of [4]

2. Generally outperform both

3. Can be used for any number of transmit and receive antennas

4. Are very simple to encode

5. Can be decoded using a variety of simple linear-algebraic techniques such as successive nulling and cancelling (V-BLAST [2], square-root V-BLAST [6]), or sphere decoding [7]

6. Are designed with the numbers of both the transmit *and* receive antennas in mind

7. Satisfy an information-theoretic optimality criterion

We briefly summarize the general structure of the LD codes. Suppose that there are 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. The transmitted signal can then be written as a $T \times M$ matrix S that governs the transmission over the M antennas during the interval. We assume that the data sequence has been broken into Q substreams and that $s_1, \ldots, s_Q$ are the complex symbols chosen from an arbitrary, say r-PSK or r-QAM, constellation. We call a rate $R = (Q/T) \log_2 r$ linear dispersion code one for which S obeys

$$S = \sum_{q=1}^{Q} (\alpha_q A_q + j\beta_q B_q), \qquad (1)$$

where the real scalars $\{\alpha_q, \beta_q\}$ are given by $s_q = \alpha_q + j\beta_q$, $q = 1, \ldots Q$. The code is completely specified by the fixed $T \times M$ complex matrices $A_1, \ldots, A_Q$ and $B_1, \ldots, B_Q$. Each individual codeword is determined by the complex scalars $\{s_1, \ldots, s_Q\}$.

Many known codes can be identified as special cases. The 2×2 code of [3]

$$S = \begin{bmatrix} s_1 & s_2 \\ -s_2^* & s_1^* \end{bmatrix}. \tag{2}$$

for example, can be obtained as a special case of (1) by taking $T = M = Q = 2$ and

$$A_1 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, A_2 = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}, B_1 = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}, B_2 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$

As another example, V-BLAST in [2] transmits individual substreams on each antenna. There is no spatial channel code structure and hence we can take $T = 1$ and $S = [s_1, s_2, \ldots, s_M]$. This is obtained in (1) by taking $Q = M$ and

$$A_q = B_q = [\underbrace{0, \ldots, 0}_{q-1}, 1, \underbrace{0, \ldots, 0}_{M-q}], \qquad q = 1, \ldots, Q.$$

The design of LD codes depends crucially on the choices of the parameters T, Q and the dispersion matrices $\{A_q, B_q\}$. To choose the $\{A_q, B_q\}$ we propose to optimize a nonlinear information-theoretic criterion: namely, the mutual information between the transmitted signals $\{\alpha_q, \beta_q\}$ and the received signal.

1.1. The Multiple-Antenna Model

Consider a narrow-band, flat-fading, multi-antenna communication system with M transmit and N receive antennas, where the channel is constant for at least T channel uses. Then the transmitted and received signals can be related by

$$X = \sqrt{\frac{\rho}{M}} SH + V, \tag{3}$$

where $X \in \mathcal{C}^{T \times N}$ is the received matrix, $S \in \mathcal{C}^{T \times M}$ is the transmitted matrix, and $V \in \mathcal{C}^{T \times N}$ is the additive noise matrix with independent $\mathcal{CN}(0, 1)$ entries. In X, S, and V, time runs vertically and space runs horizontally. $H \in \mathcal{C}^{M \times N}$ is the channel matrix of independent $\mathcal{CN}(0, 1)$ entries. The entries of the transmitted matrix S are assumed to have, on the average, unit variance entries, which implies the power constraint $\mathrm{E}\,\mathrm{tr}\,SS^* = TM$ and, along with the normalization $\sqrt{\frac{\rho}{M}}$, that ρ is the SNR at the receiver, independently of M.

H is assumed to be known to the receiver but not to the transmitter, in which case the channel capacity (often referred to as the *perfect-knowledge* capacity) is [1]

$$C(\rho, M, N) = \mathrm{E}\,\log\det\left(I_N + \frac{\rho}{M} HH^*\right). \tag{4}$$

2. INFORMATION-THEORETIC ANALYSIS OF SOME SPACE-TIME CODES

To motivate the LD codes let us begin by reviewing some existing space-time codes.

2.1. Mutual information attainable with orthogonal designs

An *orthogonal design* is introduced by Alamouti in [3] for $T = M = 2$ and has the structure given in (2). The complex scalars s_1 and s_2 are drawn from a particular (r-PSK or r-QAM) constellation, but we may simply assume that they are random variables such that $\mathrm{E}(|s_1|^2 + |s_2|^2) = 2$. One can see from (2) that S has orthogonal columns and, moreover, that any linear combination of S's also has orthogonal columns.

We may ask how much mutual information the orthogonal design structure (2) can attain? To answer this question we need to compute $C_{orth}(\rho, N)$, the mutual information between the transmitted and received signals, and compare it with $C(\rho, M, N)$, the capacity of the $M \times N$ multi-antenna system. This is done in [8], with the result

$$C_{orth}(\rho, N) = C(N \cdot \rho, M = 2 \cdot N, N = 1)$$

The above equations imply that the orthogonal design (2) can achieve full channel capacity of the $M = 2$, $N = 1$ system, but not of the $M = 2$, $N \geq 2$ system. Thus, when $N \geq 2$ we incur a loss with the structure (2), which can be substantial at high SNR.

2.2. Mutual information attainable with V-BLAST

In V-BLAST, the maximum mutual information that can be achieved is indeed the full multi-antenna channel capacity. When $N \geq M$ there exist efficient schemes for decoding the V-BLAST matrices, the best implementations of which require $O(N^3)$ computations. Nevertheless, V-BLAST suffers from two deficiencies: the simple decoding algorithms fail when $N < M$, and there is no built-in spatial coding.

3. LINEAR-DISPERSION SPACE-TIME CODES

We call a *linear-dispersion (LD) code* one for which S satisfies (1) where the A_q and B_q are *fixed* $T \times M$ complex matrices. Without loss of generality, we assume that $\alpha_1, \ldots, \alpha_Q$ and $\beta_1, \ldots, \beta_Q$ have variance $\frac{1}{2}$ and are uncorrelated. This induces the normalization

$$\sum_{q=1}^{Q} \left(\mathrm{tr}\,A_q^* A_q + \mathrm{tr}\,B_q^* B_q\right) = 2TM. \tag{5}$$

Note that in V-BLAST each signal $\{\alpha_q, \beta_q\}$ is transmitted from only one antenna during only one channel use. With

the LD codes, however, the dispersion matrices potentially transmit some combination of *each* symbol from *each* antenna at *every* channel use. This can lead to desirable coding properties.

3.1. Decoding

An important property of the LD codes (1) is their linearity in the variables $\{\alpha_q, \beta_q\}$. This leads to the following linear relation between the $\{\alpha_q, \beta_q\}$ and the received signals:

$$\underbrace{\begin{bmatrix} x_{R,1} \\ x_{I,1} \\ \vdots \\ x_{R,N} \\ x_{I,N} \end{bmatrix}}_{\triangleq x} = \sqrt{\frac{\rho}{M}} \mathcal{H} \underbrace{\begin{bmatrix} \alpha_1 \\ \beta_1 \\ \vdots \\ \alpha_Q \\ \beta_Q \end{bmatrix}}_{\triangleq s} + \underbrace{\begin{bmatrix} v_{R,1} \\ v_{I,1} \\ \vdots \\ v_{R,N} \\ v_{I,N} \end{bmatrix}}_{\triangleq v}, \quad (6)$$

where $x_{R,n}$, $x_{I,n}$, $v_{R,n}$, and $v_{I,n}$, $n = 1, \ldots, N$, denote the real and imaginary parts of the columns of X and V, respectively, and where $\mathcal{H}$ is an $2NT \times 2Q$ real *equivalent* channel matrix whose entries depend on the entries of H and the dispersion matrices $\{A_q, B_q\}$; for an explicit description of $\mathcal{H}$ see [8].

Note that the equivalent channel $\mathcal{H}$ is known to the receiver because the original channel H, and the dispersion matrices $\{A_q, B_q\}$ are all known to the receiver. When $Q \leq NT$ there are more equations than unknowns and we may therefore use any decoding technique already in place for V-BLAST to decode the signals $\{\alpha_q, \beta_q\}$.

3.2. Design of the Dispersion Codes

Although we have introduced the LD structure we have not yet specified Q or the dispersion matrices $A_1, \ldots, A_Q$ and $B_1, \ldots, B_Q$. We have the inequality $Q \leq NT$. Intuitively, the larger Q is, the higher the maximum mutual information is, and the smaller Q is, the more of a coding effect we obtain.

We are left with the question of how to design the dispersion matrices. We choose $\{A_q, B_q\}$ to maximize the mutual information between s and x in (6). This guarantees that we are taking the smallest possible mutual information penalty within the LD structure (1).

The Design Method

1. Choose $Q \leq NT$ (typically, $Q = \min(M, N) \cdot T$).

2. Choose $\{A_q, B_q\}$ that solve the optimization problem

$$\max_{A_q, B_q, q=1,\ldots Q} \frac{1}{2T} \mathrm{E} \log \det \left(I_{2NT} + \frac{\rho}{M} \mathcal{H} \mathcal{H}^t \right) \quad (7)$$

for an SNR ρ of interest, subject to one of the following constraints

(i) $\sum_{q=1}^{Q} \left(\mathrm{tr}\, A_q^* A_q + \mathrm{tr}\, B_q^* B_q \right) = 2TM$

(ii) $\mathrm{tr}\, A_q^* A_q = \mathrm{tr}\, B_q^* B_q = \frac{TM}{Q}, q = 1, \ldots Q$

(iii) $A_q^* A_q = B_q^* B_q = \frac{T}{Q} I_M, q = 1, \ldots Q$

We now make some remarks:

1. The constraints are listed in increasing order of restrictiveness. Constraint (iii) generally gives the best coding effects.

2. The solution to (7) subject to any of the constraints (i)–(iii) is highly nonunique. For example, the transformation

$$A'_q = \sum_{p=1}^{Q} \left(A_p \phi_{2p-1,2q-1} + B_p \phi_{2p,2q-1} \right)$$
$$B'_q = \sum_{p=1}^{Q} \left(A_p \phi_{2p-1,2q} + B_p \phi_{2p,2q} \right)$$

$$(8)$$

where $\Phi = \{\phi_{i,j}\}_{i,j=1}^{2Q}$ is a real orthogonal matrix, preserves the mutual information.

3. The cost function $\frac{1}{2T} \mathrm{E} \log \det \left(I_{2NT} + \frac{\rho}{M} \mathcal{H} \mathcal{H}^t \right)$ can have local maxima. Nevertheless, we have been able to solve (7) with relative ease using gradient-based methods.

4. When $N \geq M$ and $Q = MT$ one solution to (7), subject to either constraints (i) or (ii), is given by the V-BLAST matrices. However, these do not satisfy constraint (iii). A solution that satisfies (iii) for $T = M$ is

$$A'_{M(k-1)+l} = B'_{M(k-1)+l} = \frac{1}{\sqrt{M}} D^{k-1} \Pi^{l-1}, \quad (9)$$

where $k = 1, \ldots, M$, $l = 1, \ldots, M$, D is a diagonal matrix with $e^{j\frac{2\pi k}{M}}$ on its diagonals and Π is a permutation matrix.

5. The design criterion (7) *depends explicitly on the number of receive antennas N*.

4. EXAMPLES OF LD CODES AND PERFORMANCE

LD vs. OD: $M = N = 2, R = 8$

We test the performance of the 2×2 orthogonal design when $M = N = 2$ at $R = 8$ versus the LD code given by (9) for $T = 2$ and $Q = 4$. The result can be seen in Figure 1 which clearly shows the superior performance of the LD code. To achieve $R = 8$, the orthogonal design needs to choose s_1 and s_2 from a 256-QAM constellation, while the LD code can choose from a 16-QAM constellation. We note that the orthogonal design achieves 7.47 bits/channel use mutual information at $\rho = 20$ dB, while the LD code achieves the full channel capacity of 11.28 bits/channel use.
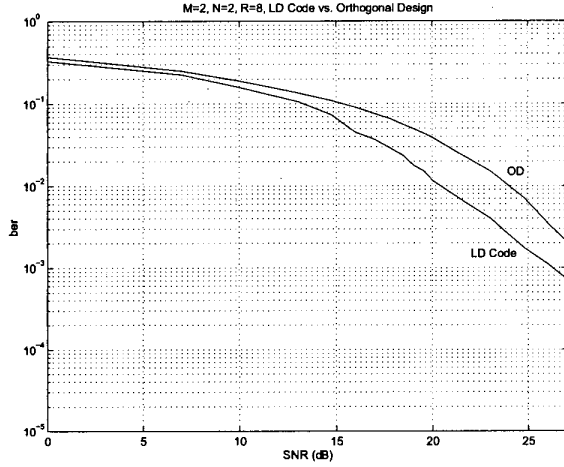
Figure 1: Bit error performance of the 2×2 orthogonal design versus the LD code given by (9) for $M = N = T = 2$ and $Q = 4$ at rate $R = 8$. The decoding is maximum likelihood.
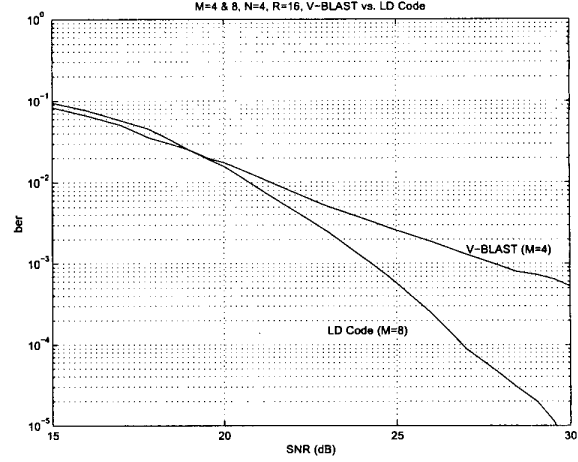
Figure 2: Bit error performance of V-BLAST for $M = 4$ and $N = 4$ and an LD code for $M = 8$ and $N = 4$ with $Q = 32$ at rate is $R = 16$ with nulling/cancelling decoding.

LD vs. V-BLAST: $M = 4, 8, N = 4, R = 16$

Figure 2 compares an LD code for $M = 8$, $N = 4$ with $M = 4$, $N = 4$ V-BLAST at rate $R = 16$. The $M = N = 4$ and $M = 8$, $N = 4$ channel capacities at 20 dB are 22.14 and 24.94 bits/channel use, respectively, whereas the optimized LD code achieves 23.10 bits/channel use with $T = 8$ and $Q = 32$.

Table 1 summarizes the mutual informations of some LD codes that we generated and the actual channel capacities at $\rho = 20$ dB. As can be observed, $C_{\mathrm{LD}}(\rho, T, M, N)$ is very close to $C(\rho, M, N)$; there is little penalty in the linear structure of the dispersion codes.

(T,M,N)	$C_{LD}(\rho, T, M, N)$	$C(\rho, M, N)$
(2,2,1)	6.28	6.28
(4,3,1)	6.25	6.41
(6,3,1)	6.28	6.41
(4,3,2)	11.63	12.14
(4,4,1)	6.34	6.47
(4,4,2)	11.99	12.49
(8,8,4)	23.10	24.94
$(T = M, M, N \geq M)$	$C(\rho, M, N)$	*

Table 1: Mutual information $C_{LD}(\rho, T, M, N)$ obtained via gradient-ascent optimization of the cost function (7), compared to the actual channel capacity $C(\rho, M, N)$.

REFERENCES

[1] I. E. Telatar, "Capacity of multi-antenna gaussian channels," *Eur. Trans. Telecom.*, vol. 10, pp. 585–595, Nov. 1999.

[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," *Electronic Letters*, vol. 17, Nov. 1998.

[3] S. M. Alamouti, "A simple transmitter diversity scheme for wireless communications," *IEEE J. Sel. Area Comm.*, pp. 1451–1458, Oct. 1998.

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

[5] B. Hassibi, A. Shokrollahi, B. Hochwald, and W. Sweldens, "Representation theory for high-rate multiple-antenna code design," *submitted to IEEE Trans. Info. Theory*, 2000. Download available at http://mars.bell-labs.com.

[6] B. Hassibi, "An efficient square-root algorithm for blast," *submitted to IEEE Trans. Sig. Proc.*, 1999. Download available at http://mars.bell-labs.com.

[7] O. Damen, A. Chkeif, and J.-C. Belfiore, "Lattice code decoder for space-time codes," *IEEE Comm. Let.*, pp. 161–163, May 2000.

[8] B. Hassibi and B. Hochwald, "High rate codes that are linear in space and time," *submitted to IEEE Trans. Info. Theory*, 2000. Download available at http://mars.bell-labs.com.