Layered Space-Time Codes for Wireless Communications Using Multiple Transmit Antennas

Da-shan Shiu and Joseph M. Kahn
University of California at Berkeley

Abstract – **Multiple-antenna systems provide very high capacity compared to single antenna systems in a Rayleigh fading environment. Space-time codes are channel codes designed to exploit this high capacity for multiple-antenna systems without requiring instantaneous channel knowledge at the transmitter. A practical concern for high data rate space-time codes is their decoding complexity. The decoding complexity with ML criterion can be prohibitively large. In this paper, we focus on layered space-time (LST) codes. Two types of LST codes, the horizontally-layered space-time (HLST) codes and the diagonally-layered space-time (DLST) codes, are presented. We analyze the performance of both types of LST codes under slow and fast fading conditions. We conclude that, in a slow fading environment, DLST codes have superior performance over HLST codes. The design criteria for DLST codes are proposed.**

I. INTRODUCTION

Recent studies have explored the ultimate limit of multiple-antenna systems from the information-theoretic point of view [1] - [3]. Consider a multiple-antenna system that has n transmitting and m receiving antennas. It is shown that, if the narrowband slow fading channel can be modeled as an $n \times m$ matrix with i.i.d. complex Gaussian random entries, the average channel capacity of such a system is approximately $\min(n, m)$-times higher than that of a single antenna system for the same overall transmitting power.

In most applications, a major obstacle to utilizing this high throughput is that the transmitter cannot have the instantaneous information about the fading channel. The transmitter thus must employ a channel code that can guarantee good performance with the majority of possible channel realizations. Such a channel code is inherently multi-spatial-dimensional and thus is called a space-time code [4] [5].

Aside from the consideration of combatting channel uncertainty, another practical consideration for space-time codes is the decoding complexity. As stated above, the channel capacity of a multiple-antenna system is approximately proportional to $\min(n, m)$. This means that each channel usage on average can convey proportional to $\min(n, m)$ bits of information. The complexity of decoding such a high data-rate channel code using the maximal-likelihood (ML) criterion can be prohibitively high even if $\min(n, m)$ is just moderately large; thus, space-time codes that admit high performance, low complexity suboptimal decoding algorithms are desirable.

The layered space-time (LST) architecture proposed by Foschini in [6] is a framework of processing space-time signals. An LST code is a channel code that is designed and processed according to the LST architecture. An LST code is constructed by assembling 1-D constituent codes. With the use of interference suppression and interference cancellation at the receiver, these constituent codes can be separated and then decoded using conventional decoding algorithm developed for 1-D constituent codes, leading to a much lower decoding complexity compared to ML decoding. Other possible low-complexity decoding techniques include sequential decoding (SD) [7] and multistage decoding [4] [8].

In this paper, we analyze the performance of both horizontally-layered space-time (HLST) codes and diagonally-layered space-time (DLST) codes. We will show that DLST codes can achieve better performance compared to HLST codes. From our analysis, we propose the design criteria, i.e., the criteria to choose the constituent code, for DLST codes based on the truncated multidimensional effective code length (TMEL) and the truncated multidimensional product distance (TMPD) of the constituent code.

The remainder of this paper is organized as follows. In Section II, the background of multiple-antenna systems and space-time codes is reviewed. In Section III, we introduce the LST architecture. In Section IV, the performances of HLST and DLST codes are analyzed. The analysis leads to the design criteria for DLST codes. We also present an example DLST code. Concluding remarks can be found in Section V.

II. BACKGROUND

In this paper, we focus on single-user to single-user communication using multiple antennas at both ends over narrowband flat-fading channels. We refer to a multiple-antenna system in which the transmitter has n transmitting antennas and the receiver has m receiving antennas as an (n, m) system. A general space-time code can be described as follows. The encoder first applies the space-time code to the input information bits to generate an n-row (possibly semi-infinite) matrix $\boldsymbol{C}$. The matrix $\boldsymbol{C}$ represents the signal that is to be transmitted by the transmitter. Specifically, the kth row, τth column element of $\boldsymbol{C}$, denoted by c_τ^k, represents the signal to be transmitted by antenna k at time slot τ. We emphasize that there is no mechanism, such as time-, frequency-, or code-division multiplexing, employed to ensure that the signals transmitted by different transmitting antennas are orthogonal upon reception by the receiver. The signal received by the receiving antenna l during the time slot τ is denoted by r_τ^l. This received signal r_τ^l contains a superposition of transmitted signals c_τ^k, $k = 0, 1, \ldots, n-1$, and an AWGN component v_τ^l.

For a narrowband flat-fading channel, the gain connecting transmitting antenna k and receiving antenna l at time τ can be denoted by a complex number $h_\tau^{l \leftarrow k}$. Define the vectors[1] $\boldsymbol{c}_\tau = (c_\tau^1\ c_\tau^2\ \ldots c_\tau^n)'$, $\boldsymbol{r}_\tau = (r_\tau^1\ r_\tau^2\ \ldots r_\tau^m)'$, $\boldsymbol{v}_\tau = (v_\tau^1\ v_\tau^2\ \ldots v_\tau^m)'$, and the channel matrix H_τ, $(H_k^l)_\tau = h_\tau^{l \leftarrow k}$. The discrete-time input-output relation of an (n, m) multiple-antenna system over a narrowband flat-fading channel can be written in the following vector notation:

$$\boldsymbol{r}_\tau = H_\tau \boldsymbol{c}_\tau + \boldsymbol{v}_\tau . \qquad (1)$$

The following terminology is used in this paper. The matrix $\boldsymbol{C}$, a coded matrix output of the transmitter encoder, is referred to as a space-time *codeword matrix*. A space-time codeword matrix can be thought of as a serial concatenation of *n-tuples*, and an n-tuple is composed of *n symbols*. Note that the first row of the matrix $\boldsymbol{C}$ is indexed as row zero, not row one. To facilitate the comparison between multiple transmit-antenna systems using space-time codes and single transmit-antenna systems using conventional 1-D channel codes at equal average transmit powers (total over all transmit antennas), the average energy of an n-tuple is E, regardless of the spatial dimensionality n.

III. The Layered Space-Time Architecture

A. Encoding

The encoding process is illustrated in Fig. 1. In Fig. 1(a), the input information bit sequence is first demultiplexed into n subsequences, and each subsequence is subsequently encoded by a 1-D encoder. These 1-D channel codes are referred to as the constituent codes (CC). The output of the constituent coder k is a sequence of symbols s_τ^k, $\tau = 0, 1, \ldots$. In the LST architecture, the multi-spatial dimensional codeword matrix $\boldsymbol{C}$ is constructed by assigning these symbols to the slots of $\boldsymbol{C}$ in a systematic fashion with the goal of reducing the receiver complexity.

One intuitive assignment rule is to simply place s_τ^k at the kth row, τth column slot of $\boldsymbol{C}$. With this assignment rule, the output coded symbol from constituent encoder k is always transmitted using the transmit antenna k. This is illustrated in Fig. 1(b). Under this assignment rule, the space-time codeword matrix has an obvious horizontally layered structure. Therefore, it is referred to as the horizontally layered space-time (HLST) architecture. HLST was originally proposed by Foschini in [6]. Another assignment rule, also proposed in [6], is the diagonally layered space-time (DLST) architecture. In DLST, instead of always sending the output symbols from a constituent coder to a particular transmit antenna, they are fed to the n transmitting antennas in turn. DLST is illustrated in Fig. 1(c). Using our codeword matrix notation, to construct the DLST codeword matrix $\boldsymbol{C}$, the output symbols from constituent coder 0 are used to fill the leftmost NW-SE diagonal of $\boldsymbol{C}$, and the output symbols from constituent coder 1 are used to fill the next diagonal, and so on. The layered structure of both HLST and DLST codeword matrices are shown in Fig. 2.

If the data rate of the constituent code maintains constant regardless of n, the data rate of an LST code is obviously proportional to n.

B. Decoding

To decode an LST code, the received signal is processed along both the spatial and temporal dimensions. Here, we introduce the spatial signal processing in the LST architecture. Focus on a given instance in time, say τ. The transmitted n-tuple is $\boldsymbol{c}_\tau$, and the received m-tuple is $\boldsymbol{r}_\tau = H_\tau \boldsymbol{c}_\tau + \boldsymbol{v}_\tau$. The received signal $\boldsymbol{r}_\tau$ is a superposition of transmitted coded symbols scaled by the channel gain and corrupted by AWGN. The task here is to determine the values of the n components

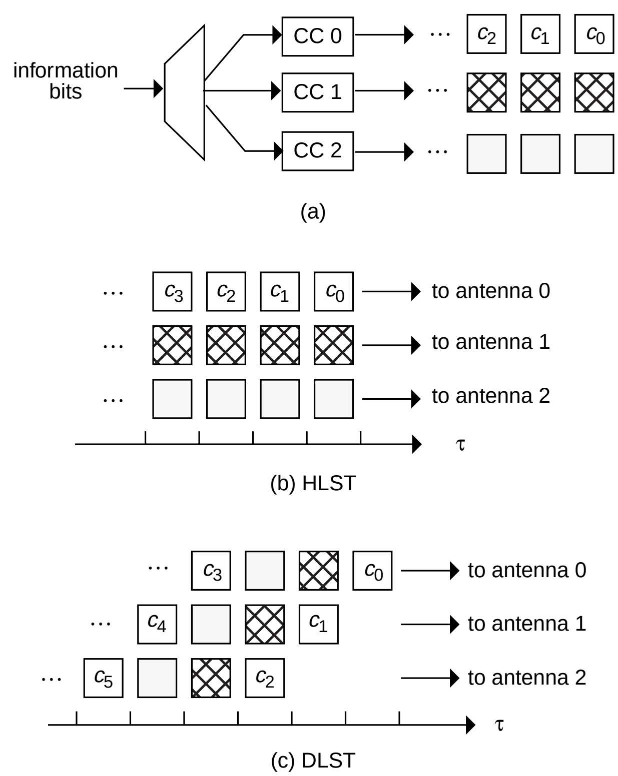

Fig. 1. LST code encoding process. Here, $n = 3$. Each square represents a symbol. (a) The incoming information bit sequence is first demultiplexed into n subsequences. Each subsequence is then encoded using a constituent code. (b) In HLST, the coded symbols from constituent encoder k are transmitted by antenna k. (c) In DLST, the coded symbols from a constituent encoder are transmitted by the n transmitting antennas in turn.

[1] The transpose and conjugate transpose of x are denoted by x', and $x^\dagger$, respectively.

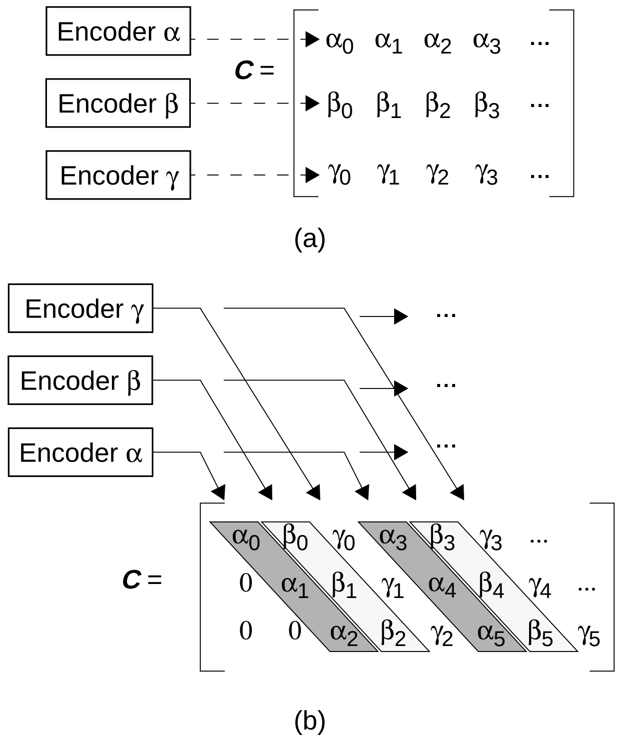

Fig. 2. The layered structure of an LST codeword matrix. (a) HLST. (b) DLST.

of $\boldsymbol{c}_\tau$, i.e. c_τ^0, c_τ^1, …, c_τ^{n-1}, with the only available information being $\boldsymbol{r}_\tau$ and H_τ.

Under the LST architecture, the decisions on the values of these n components are made sequentially according to a predetermined order. In this paper, without loss of generality, given a fixed τ, c_τ^k is decided in descending order of k. Consider the symbol c_τ^{n-1}, which is to be detected first. An interference suppression operation is employed to extract a decision variable about c_τ^{n-1} from the received signal $\boldsymbol{r}_\tau$. This decision variable, denoted by z_τ^{n-1}, contains a low level of interference from other transmitted symbols. The decision on c_τ^{n-1} is then made based on z_τ^{n-1}. Making use of the decision on c_τ^{n-1}, the receiver can modify the received signal $\boldsymbol{r}_\tau$ by removing the contribution of c_τ^{n-1} to it. This modifying operation is referred to as interference cancellation. The process of interference suppression, symbol value decision, and interference cancellation is repeated for the remaining symbols c_τ^{n-2}, c_τ^{n-3}, …, c_τ^0.

Suppose that the interference from other transmitted symbols is to be completely suppressed using linear operation. Let the QR decomposition of H_τ be $H_\tau = (U_R)_\tau R_\tau$, where $(U_R)_\tau$ is a unitary matrix and R_τ is an upper triangular matrix. Left-multiply the received signal $\boldsymbol{r}_\tau$ by $(U_R)_\tau$,

$$\boldsymbol{y}_\tau = (U_R)_\tau \boldsymbol{r}_\tau = R_\tau \boldsymbol{c}_\tau + \boldsymbol{v}_\tau', \tag{2}$$

where $\boldsymbol{v}_\tau' = (U_R)_\tau \boldsymbol{v}_\tau$ is an m-tuple of i.i.d. AWGN noise component. Because R_τ is upper triangular,

$$y_\tau^k = (R_k^k)_\tau c_\tau^k + v_\tau'^k + \{\text{constribution from } c_\tau^{k+1}, c_\tau^{k+2}, \ldots, c_\tau^{n_T}\}. \tag{3}$$

Note that in (3) only the interferences from c^l, $l < k$, are suppressed in y^k. The interference term in (3) is canceled by using the available decisions $\hat{c}_\tau^{k+1}, \hat{c}_\tau^{k+2}, \ldots, \hat{c}_\tau^{n-1}$. Assuming that these decisions are all correct, the decision variable z_τ^k is

$$z_\tau^k = (R_k^k)_\tau c_\tau^k + v_\tau'^k, \; k = 0, 1, \ldots, n-1. \tag{4}$$

The complexity of the spatial processing described above is $O(n^2 + nm)$ per transmitted n-tuple. Other interference suppression criteria, such as the linear least square criterion, can also be employed instead [9].

IV. Layered Space-Time Codes

A. HLST

To decode an HLST codeword matrix $\boldsymbol{C}$, the receiver first extracts the decision variables for the symbols of the bottom-most row of $\boldsymbol{C}$. The resulting decision variable sequence, $\{z_\tau^{n-1}\}$, $\tau = 0, 1, \ldots$, is then passed to a conventional 1-D decoder of the corresponding constituent code to produce the decisions on the symbols of this row. The receiver uses these decisions to modify the received signal sequence $\{\boldsymbol{r}_\tau\}$, and then proceeds to decode row $n-2$, $n-3$, and so on. In short, the HLST codeword matrix $\boldsymbol{C}$ is decoded row by row, or layer by layer, from bottom to top.

Consider the kth row of an HLST codeword matrix $\boldsymbol{C}$. Let $\{c_\tau^k\}$ denote the actual transmitted symbol sequence on this row, and $\{e_\tau^k\}$ denote a distinct possible transmitted symbol sequence. Conditioned on the channel realization $H_\tau = \{H_0, H_1, \ldots\}$, the probability that the likelihood of transmitting $\{e_\tau^k\}$ is higher than $\{c_\tau^k\}$ is

$$\begin{aligned} Prob(\boldsymbol{c}^k \to \boldsymbol{e}^k | H_\tau) &= Q\left(\sqrt{\frac{E}{2N_0}\sum_\tau \left|(R_k^k)_\tau\right|^2 \left|c_\tau^k - e_\tau^k\right|^2}\right) \\ &\le \exp\left\{-\frac{E}{4N_0}\sum_\tau \left|(R_k^k)_\tau\right|^2 \left|c_\tau^k - e_\tau^k\right|^2\right\}, \end{aligned} \tag{5}$$

where the matrix R_τ comes from the QR decomposition of H_τ, i.e., $H_\tau = U_\tau R_\tau$, and $Q(x) = (2\pi)^{-1/2}\int_x^\infty \exp(-x^2/2)dx$. In (5), we make use of the Chernoff bound for the Q function, $Q(x) \le \exp(-x^2/2)$. $Prob(\boldsymbol{c}^k \to \boldsymbol{e}^k | H_\tau)$ is the conditional pairwise error probability between $\{c_\tau^k\}$ and $\{e_\tau^k\}$. The average pairwise error probability can thus be upper-bounded by taking the expected value of the right side of (5) over the distribution of $\left|(R_k^k)_\tau\right|^2$, which is a chi-squared distribution with $2(m-k)$ degrees of freedom [10]. In a fast fading environment, the $\left|(R_k^k)_\tau\right|^2$ are i.i.d. for distinct τ. The average pairwise error probability, $Prob(\boldsymbol{c}^k \to \boldsymbol{e}^k)$, can be upper bounded by:

$$
\begin{aligned}
&Prob(\boldsymbol{c}^k \to \boldsymbol{e}^k) \\
&\le \prod_{\tau \in \eta(\boldsymbol{c}^k, \boldsymbol{e}^k)} \boldsymbol{E}\left[\exp\left\{-\frac{E}{4N_0}\left|(R_k^k)_\tau\right|^2\left|c_\tau^k - e_\tau^k\right|^2\right\}\right] \\
&= \prod_{\tau \in \eta(\boldsymbol{c}^k, \boldsymbol{e}^k)} \left(1 + \left|c_\tau^k - e_\tau^k\right|^2 \frac{E}{4N_0}\right)^{-(m-k)},
\end{aligned}
\tag{6}
$$

where $\eta(\boldsymbol{c}^k, \boldsymbol{e}^k) = \{\tau | c_\tau^k \ne e_\tau^k\}$. In a slow fading environment, $\left|(R_k^k)_\tau\right|^2 = \left|R_k^k\right|^2$ for all τ, and $Prob(\boldsymbol{c}^k \to \boldsymbol{e}^k)$ can be upper bounded by:

$$
\begin{aligned}
Prob(\boldsymbol{c}^k \to \boldsymbol{e}^k) &\le \boldsymbol{E}\left[1 + \frac{E}{4N_0}\left|R_k^k\right|^2 \sum_\tau \left|c_\tau^k - e_\tau^k\right|^2\right]^{-(m-k)} \\
&= \left(1 + \frac{E}{4N_0}\left|\boldsymbol{c} - \boldsymbol{e}\right|^2\right)^{-(m-k)}.
\end{aligned}
\tag{7}
$$

B. DLST

A DLST codeword matrix is decoded diagonal by diagonal. To illustrate this, consider the example DLST codeword matrix in Fig. 2. The receiver first generates the decision variables for the symbols of the first diagonal of $\boldsymbol{C}$, namely α_0, α_1, and α_2. Based on the decision variables, this diagonal is decoded, and the decision is then fed back to remove the contribution of this diagonal to the received signal. The receiver then continues to decode the next diagonal, and so on.

Here we consider the probability of a diagonal decision error. Consider the leftmost NW-SE diagonal of a DLST codeword matrix. On this diagonal, the transmitted symbols are c_τ^τ, $\tau = 0, 1, \ldots, n-1$. The probability that, under the DLST decoding algorithm, the likelihood of a distinct diagonal $\boldsymbol{e} = \{e_0^0\ e_1^1\ \ldots\ e_{n-1}^{n-1}\}$ is higher than that of the transmitted diagonal $\boldsymbol{c} = \{c_0^0\ c_1^1\ \ldots\ c_{n-1}^{n-1}\}$, conditioned on the channel realization $H_\tau = \{H_0, H_1, \ldots\}$, can be derived by applying (3) noting that the interference term in (3) is zero for this diagonal:

$$
\begin{aligned}
Prob(\boldsymbol{c} \to \boldsymbol{e} | H_\tau) &= Q\left(\sqrt{\frac{E}{2N_0}\sum_{\tau=0}^{n-1}\left|(R_\tau^\tau)_\tau\right|^2\left|c_\tau^\tau - e_\tau^\tau\right|^2}\right) \\
&\le \exp\left\{-\frac{E}{4N_0}\sum_{\tau=0}^{n-1}\left|(R_\tau^\tau)_\tau\right|^2\left|c_\tau^\tau - e_\tau^\tau\right|^2\right\}.
\end{aligned}
\tag{8}
$$

Equation (8) applies in both fast and slow fading environments because the $\left|(R_\tau^\tau)_\tau\right|^2$ are i.i.d. for $\tau = 0, 1, \ldots, n-1$.

The upper bound of the average pairwise error probability is again obtained by taking the expected value of the right-hand side of (8). When the SNR is high,

$$
\begin{aligned}
Prob(\boldsymbol{c} \to \boldsymbol{e}) &\le \prod_{\tau \in \eta(\boldsymbol{c}, \boldsymbol{e})} \left(1 + \left|c_\tau^\tau - e_\tau^\tau\right|^2 \frac{E}{4N_0}\right)^{-(m-\tau)} \\
&\approx \left\{\prod_{\tau \in \eta(\boldsymbol{c}, \boldsymbol{e})} (\left|c_\tau^\tau - e_\tau^\tau\right|^2)^{-(m-\tau)}\right\}\left(\frac{E}{4N_0}\right)^{-\sum_{\tau \in \eta(\boldsymbol{c}, \boldsymbol{e})} m-\tau},
\end{aligned}
\tag{9}
$$

where $\eta(\boldsymbol{c}, \boldsymbol{e}) = \{\tau | c_\tau^\tau \ne e_\tau^\tau\}$. When the SNR is low, i.e. $\left|c_\tau^\tau - e_\tau^\tau\right|^2 E/4N_0 \ll 1$ for all τ, using the approximation $(1 + mx)^{-1} \approx (1 + x)^{-m}$ for small mx, (9) can be approximated by

$$
\begin{aligned}
Prob(\boldsymbol{c} \to \boldsymbol{e}) &\le \prod_{\tau \in \eta(\boldsymbol{c}, \boldsymbol{e})} \left[\left(1 + \varepsilon\frac{E}{4N_0}\right)^{\varepsilon^{-1}}\right]^{-\{\left|c_\tau^\tau - e_\tau^\tau\right|^2(m-\tau)\}} \\
&= \left[\left(1 + \varepsilon\frac{E}{4N_0}\right)^{\varepsilon^{-1}}\right]^{-\left\{\sum_{\tau \in \eta(\boldsymbol{c}, \boldsymbol{e})} \left|c_\tau^\tau - e_\tau^\tau\right|^2(m-\tau)\right\}},
\end{aligned}
\tag{10}
$$

where ε is an arbitrarily small positive number.

C. Comparison

By comparing (7) and (9), we conclude that the performance of HLST codes in slow fading environments is inferior to that of DLST codes. For an HLST code, the average pairwise error probability of the bottommost row is inversely proportional to the $(m-n+1)$th power of SNR. In contrast, in DLST, the average pairwise diagonal error probability between two diagonals $\boldsymbol{c}$ and $\boldsymbol{e}$ is inversely proportional to the $(\sum_{\tau \in \eta(\boldsymbol{c}, \boldsymbol{e})} m - \tau)$th power of SNR. Therefore, if constituent codes of equivalent data rate and complexity are deployed, the error probability of a DLST code in a slow fading environment can be much lower than that of an HLST code.

D. Design Criteria for DLST Codes

Define the truncated multi-dimensional effective length (TMEL) and the truncated multi-dimensional product distance (TMPD) between two distinct diagonals $\boldsymbol{c}$ and $\boldsymbol{e}$ as

$$
\text{TMEL} \sum_{\tau \in \eta(\boldsymbol{c}, \boldsymbol{e})} m - \tau \text{ and} \tag{11a}
$$

$$
\text{TMPD} \equiv \prod_{\tau \in \eta(\boldsymbol{c}, \boldsymbol{e})} \left|c_\tau^\tau - e_\tau^\tau\right|^{2(m-\tau)}. \tag{11b}
$$

At high SNR, the pairwise error probability between $\boldsymbol{c}$ and $\boldsymbol{e}$ is approximated by $Prob(\boldsymbol{c} \to \boldsymbol{e}) \approx (\text{TMPD})^{-1}(E/4N_0)^{-\text{TMEL}}$. The code design criterion is to maximize the minimum value of $Prob(\boldsymbol{c} \to \boldsymbol{e}) \approx (\text{TMPD})^{-1}(E/4N_0)^{-\text{TMEL}}$ over all pairs of distinct diagonals. If the exact operating SNR is not known but can be assumed to be reasonably high, an approximate design criterion is to maximize the minimum two-tuple (TMEL, TMPD) in dictionary order.

At low SNR, the pairwise error probability is approximated by (10). We define the exponent $\sum_{\tau \in \eta(\boldsymbol{c}, \boldsymbol{e})} \left|c_\tau - e_\tau\right|^2(m-\tau)$ to be the truncated multi-dimensional Euclidean distance

(TMED) between ***c*** and ***e***. The code design criterion at low SNR is to maximize the minimum TMED between any pair of distinct diagonals.

E. Example

In this example, a (7, 3) RS code is used as a constituent code of a 7-D DLST code. The (7, 3) RS encoder maps three 8-ary input digits into seven 8-ary output digits. Each 8-ary digit selects a point (symbol) on the 8-PSK constellation according to the Gray code mapping, and these seven constellation points are used to fill the slots of a diagonal of the 7-D DLST code. The minimum TMEL of this code given $n = m = 7$ is 15.

Monte-Carlo simulations are performed to obtain the performance of this DLST code in slow fading environments given $n = m = 7$. Fig. 3 shows the average error probability assuming perfect decision feedback. Fig. 3 shows that, with an average SNR of 12 dB, it is possible to use this simple DLST code to achieve a data rate of 9 bits/s/Hz with an average diagonal detection error probability lower than 10^{-4}. ■

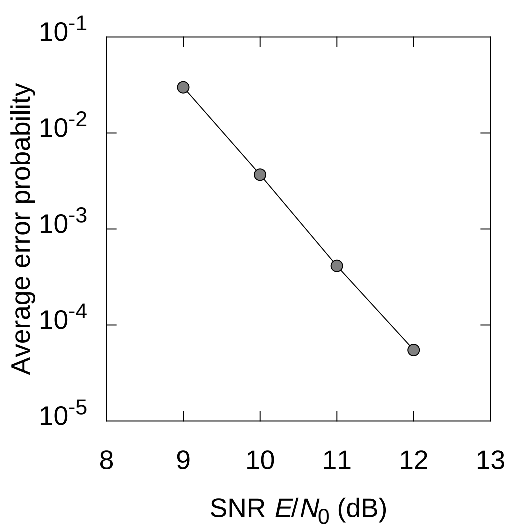

Fig. 3. Simulation results of the error probability of a 7-D DLST code using the (7, 3) RS as its constituent code in slow fading environments. Here, $n = m = 7$.

V. Summary

In this paper we considered layered space-time codes. We showed that, if the wireless channel is i.i.d. Rayleigh fading, an (n, m) multiple-antenna system employing an LST code can achieve a throughput $\min(n, m)$ times higher than that of a single-antenna system for a given overall transmit power limit.

LST codes have two important characteristics. First, the transmitter is not required to have the instantaneous channel information to employ an LST code. Secondly, LST codeword matrices are constructed from one-spatial-dimensional constituent codewords. The decoding complexity of an LST code is only quadratic in the number of antennas, making it suitable for systems that have a large number of antennas. Furthermore, the existing technology of 1-D codec can be effectively leveraged.

We analyzed the performance of both HLST and DLST codes. Our result indicates that in a slow fading environment, DLST codes have superior performance compared to HLST codes. We proposed the design criteria for DLST codes based on the following parameters of the constituent code: TMEL, TMPD, and TMED.

VI. References

[1] E. Telatar, “Capacity of Multi-antenna Gaussian Channels”, AT&T-Bell Labs Internal Tech. Memo., June 1995.

[2] G. J. Foschini and M. J. Gans, “On Limits of Wireless Communication in a Fading Environment When Using Multiple Antennas”, accepted for publication in *Wireless Personal Communications*.

[3] G. J. Foschini, R. A. Valenzuela, “Initial Estimation of Communication Efficiency of Indoor Wireless Channel”, *Wireless Networks 3*, pp. 141-154, 1997.

[4] V. Tarokh, N. Seshadri, A.R. Calderbank, “Space-Time Codes for High Data Rate Wireless Communication: Performance Criterion and Code Construction”, *IEEE Trans. Inform. Theory*, vol. 44, no. 2, March 1998.

[5] J. Guey, M.P. Fitz, M.R. Bell and W. Kuo, “Signal Design for transmitter diversity wireless communication systems over Rayleigh fading channels”, in *Proc. IEEE VTC’96*, pp. 136-40.

[6] G. J. Foschini, “Layered Space-Time Architecture For Wireless Communication in a Fading Environment When Using Multi-Element Antennas”, *Bell Labs Technical Journal*, Autumn 1996.

[7] S. Lin, D. J. Costello, *Error Control Coding: Fundamentals and Applications*, Prentice Hall, New Jersey, 1983.

[8] A.J. Calderbank, “Multilevel Codes and Multistage Decoding”, *IEEE Trans. Commun.*, vol. 37, no. 3, pp.222-29, March 1993.

[9] D. Tse, “Capacity Scaling in Multi-Antenna Systems”, in preparation.

[10] R. J. Muirhead, *Aspects of Multivariate Statistical Theory*, John Wiley & Sons, New York, 1982.