High-Rate Layered Space-Time Transmissions based on Constellation-Rotation*

Yan Xin, Zhiqiang Liu, and Georgios B. Giannakis
Dept. of ECE, Univ. of Minnesota; 200 Union Str. SE, Minneapolis, MN 55455, USA

***Abstract*— Recent theoretical and experimental studies have shown that with affordable complexity, layered space-time (LST) transmissions can attain very high spectral efficiency in a rich-scattering environment. In this paper, we propose a novel high rate linearly precoded LST system, which allows for any number of transmit and receive antennas, and offers flexibility in trading performance with bandwidth efficiency and decoding complexity. Even with sub-optimum decoding, the system enjoys considerable transmit diversity gains. Its superior performance over existing uncoded V-BLAST and Linear Dispersion (LD) codes is confirmed by simulations.**

I. INTRODUCTION

Most existing multi-antenna systems focus either on high-performance, or, on high-rate. Performance-oriented systems optimize the Bit Error Rate (BER) by enhancing diversity and (possibly coding) gains. With improved performance, higher transmission rate is usually achieved by adopting a larger constellation. Typical performance-oriented systems rely on Space-Time (ST) block codes [9], or, ST trellis codes [8]. On the other hand, rate-oriented systems focus on boosting transmission rates, while improving diversity and coding gains indirectly by incorporating channel coding, or, by resorting to receive-diversity. Typical examples are the Vertical Bell-Labs Layered ST (V-BLAST) system [3], and the ST system using Linear Dispersion (LD) codes [4]. Because both classes of multi-antenna systems cope with BER performance and transmission rate in a disjoint way, these systems are usually unable to strike the best performance-rate tradeoff, and do not offer sufficient design flexibility.

Incorporating non-redundant ST Constellation Rotation (CR) [10] into a carefully designed LST structure, we propose in this paper a novel system that we term LST-CR, to deal with the performance-rate tradeoff in a direct and effective way. The LST-CR system has the following features:

- it subsumes V-BLAST and ST-CR as special cases;
- it enables high transmission rate without necessarily enlarging the constellation size;
- it remains operational with any number of transmit and receive antennas;
- it offers flexibility to strike arbitrary performance-rate tradeoffs by tuning a single system parameter; and
- it has relatively low decoding complexity.

II. SYSTEM MODEL

Consider a multi-antenna communication system equipped with M transmit- and N receive- antennas, signaling through flat fading channels. At the transmitter, information symbols $s_i \in \mathcal{A}_s$ belonging to the constellation set $\mathcal{A}_s$ are first parsed into $Q \times 1$ blocks, $\mathbf{s} := [s_1, \ldots, s_Q]^T$, which are then mapped to an $M \times K$ Space-Time codematrix $\mathbf{C}$, whose (m,k)th entry $[\mathbf{C}]_{mk} = c_{mk}$ is transmitted through the mth transmit antenna during the kth time slot.

Suppose that perfect synchronization has been acquired. After receive-filtering and symbol-rate sampling, the kth received sample by the nth receive antenna is given by

$$y_{nk} = \sum_{m=1}^{M} h_{n,m} c_{mk} + w_{nk},\ n \in [1, N],\ k \in [1, K], \quad (1)$$

where $h_{n,m}$ is the channel coefficient between the mth transmit and nth receive antenna. We further assume that:

- **(A1)** $h_{n,m}$'s are i.i.d. zero mean complex Gaussian with variance 0.5 per dimension;
- **(A2)** $h_{n,m}$'s are known to the receiver (but not to the transmitter) and are time-invariant over K time slots;
- **(A3)** the additive noise w_{nk} is circularly complex Gaussian distributed with zero mean and variance N_0.

Denote by $\mathbf{Y}$ the $N \times K$ received sample matrix with (n,k)th entry $[\mathbf{Y}]_{nk} = y_{nk}$; $\mathbf{H}$ the $N \times M$ channel matrix with $[\mathbf{H}]_{nm} = h_{n,m}$; and, $\mathbf{W}$ the $N \times M$ noise matrix with $[\mathbf{W}]_{nk} = w_{nk}$. Under (A2), it follows from (1) that

$$\mathbf{Y} = \mathbf{HC} + \mathbf{W}. \quad (2)$$

III. LST-CR ENCODING

As a motivating observation for LST-CR, we first notice that the LST structure in V-BLAST enables high rate transmissions with reduced decoding complexity. On the other hand, constellation rotation improves performance considerably without sacrificing transmission rate. The basic idea behind our LST-CR scheme is to incorporate constellation rotation into a judiciously designed LST structure to coin a system which is capable of supporting high-rate transmissions with good performance and reduced decoding complexity.

As depicted in Fig. 1, the ST mapping is carried out in three successive steps; namely, demultiplexing, layer precoding, and ST permuting. Before we describe these steps, let us choose $M \geq K$, and select $Q = LK$, where $L := M - K + 1$. The function of the demultiplexing is to divide the information block $\mathbf{s}$ into L information sub-blocks, $\mathbf{s}^{(l)} := [s_1^{(l)}, \cdots, s_K^{(l)}]^T$, $l = 1, \ldots, L$. The function of the layer precoding is to perform constellation rotation on each $\mathbf{s}^{(l)}$ to yield $\mathbf{c}^{(l)} := \boldsymbol{\Theta}\mathbf{s}^{(l)}$, where $\boldsymbol{\Theta}$ is a $K \times K$ matrix that will be designed later. Following the BLAST terminology, we will call $\mathbf{c}^{(l)}$ the lth layer, for convenience. After layer precoding, the ST permutator takes the L layers $\{\mathbf{c}^{(l)}\}_{l=1}^{L}$ as input to

*Supported by the NSF Wireless Initiative grant no. 99-79443 and the ARL/CTA grant no. DAAD19-01-2-011

form the $M \times K$ codematrix:

$$\mathbf{C} := \begin{bmatrix} c_1^{(1)} & 0 & \dots & 0 \\ c_1^{(2)} & c_2^{(1)} & \ddots & \vdots \\ \vdots & c_2^{(2)} & \ddots & 0 \\ c_1^{(L)} & \vdots & \ddots & c_K^{(1)} \\ 0 & c_2^{(L)} & \vdots & c_K^{(2)} \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \dots & 0 & c_K^{(L)} \end{bmatrix}_{M\times K}, \qquad (3)$$

where, $c_k^{(l)}$ is the kth entry of $\mathbf{c}^{(l)}$. It follows readily that $c_{mk} = c_k^{(m-k+1)} = \boldsymbol{\theta}_k^T \mathbf{s}^{(m-k+1)}$ for $m \in [k, M-K+k]$, and 0 otherwise, where $\boldsymbol{\theta}_k^T$ stands for the kth row of the CR precoder $\boldsymbol{\Theta}$. Also recall that $c_k^{(m-k+1)}$ is transmitted through the mth transmit-antenna at the kth time slot.

Because Q information symbols are transmitted through M transmit-antennas during K time slots, we define the transmission efficiency as $\eta := Q/(MK) = (M-K+1)/M$, and the transmission rate as $R = (M-K+1)\log_2|\mathcal{A}_s|$ bps/Hz, respectively. In LST-CR, K is an important parameter under our control. Choosing different values of K offers flexibility to tradeoff among performance, transmission rate, and decoding complexity. To appreciate this flexibility of LST-CR, we present two special cases:

Case 1 ($K = 1$ and $\boldsymbol{\Theta} = 1$): Here, we have $L = M$, and $c_{mk} = s_k^{(m)} \neq 0 \; \forall \; m, k$. Thus, our LST-CR reduces to V-BLAST, which is well known to enjoy full transmission rate ($\eta = 1$), but suffers from poor performance when $M = N$;

Case 2 ($K = M$): In this case, we have $L = 1$ and $c_{mk} = 0, \; \forall \; m \neq k$, and our LST-CR specializes to ST-CR, which achieves full diversity gain MN with transmission efficiency $1/M$ [10, 11].

Cases 1 and 2 correspond to two extremes among different performance-rate tradeoffs. By choosing $1 < K < M$, LST-CR is capable of offering great flexibility to meet any desirable performance-rate requirement. To make this possible, we have to explore decoding options in LST-CR.

IV. Layer Decoding

The layered encoding of LST-CR implies that decoding can be performed in a layer by a layer fashion as well. In this section, we will focus on sub-optimum decoding, but in order to understand the motivation behind practical (albeit sub-optimum) layer decoding, we will start with the optimal ML decoding scheme.

A. *Optimum Decoding*

Because $\mathbf{W}$ is Gaussian zero-mean white, the ML estimate of $\mathbf{C}$ for a given $\mathbf{H}$ is [c.f. (2)]

$$\hat{\mathbf{C}}_{\text{opt}} = \arg\min_{\mathbf{C}\in\mathcal{C}} \text{Tr}[(\mathbf{Y}-\mathbf{HC})(\mathbf{Y}-\mathbf{HC})^{\mathcal{H}}],$$

where $\mathcal{C}$ denotes the set of all possible codematrices $\mathbf{C}$, and its cardinality is $|\mathcal{C}| = |\mathcal{A}_s|^{K(M-K+1)} = 2^{RK}$. In practice, optimal ML decoding is infeasible especially for high rate transmissions since its decoding complexity grows exponentially in R. This motivates sub-optimum LST decoding alternatives.

B. *Sub-optimum LST-CR Decoding*

Similar to V-BLAST, our sub-optimum LST decoder will include three steps per iteration: interference nulling, symbol decoding, and interference cancelling. The nulling/cancelling iteration is an efficient decoding option, but its performance suffers from error propagation. To alleviate this problem we will rely on two cures: layer ordering so that more reliable layers are decoded first, and near-optimum sphere-decoding (SD) so that the number of errors is minimized. Suppose that we are to decode the information sub-block $\mathbf{s}^{(l)}$, after the first $(l-1)$ layers have been already decoded.

(S1) *Nulling*: In this step, the interference from layers $l+1$ to L are nulled out from the kth received snapshot $\bar{\mathbf{y}}_k^{(l)}$. Defining $\hat{c}_k^{(l')}$ as the estimate of $c_k^{(l')}$ with error $e_k^{(l')} := c_k^{(l')} - \hat{c}_k^{(l')}$ for $l' \in [1, l-1]$, we have

$$\bar{\mathbf{y}}_k^{(l)} := \underbrace{\begin{bmatrix} h_{1,l+k-1} & \dots & h_{1,M-K+k} \\ \vdots & \vdots & \vdots \\ h_{N,l+k-1} & \dots & h_{N,M-K+k} \end{bmatrix}}_{:=\mathbf{H}_k^{(l)}} \underbrace{\begin{bmatrix} c_k^{(l)} \\ \vdots \\ c_k^{(L)} \end{bmatrix}}_{\mathbf{c}_k^{(l)}} + \underbrace{\begin{bmatrix} h_{1,k} & \dots & h_{1,l+k-2} \\ \vdots & \vdots & \vdots \\ h_{N,k} & \dots & h_{N,l+k-2} \end{bmatrix} \begin{bmatrix} e_k^{(1)} \\ \vdots \\ e_k^{(l-1)} \end{bmatrix} + \begin{bmatrix} w_{1k} \\ \vdots \\ w_{Nk} \end{bmatrix}}_{:=\bar{\mathbf{w}}_k^{(l)}}, \qquad (4)$$

Let us partition $\mathbf{H}_k^{(l)}$ as: $\mathbf{H}_k^{(l)} := [\mathbf{h}_k^{(l)} \; : \; \bar{\mathbf{H}}_k^{(l)}]$, where $\mathbf{h}_k^{(l)}$ is the first column of $\mathbf{H}_k^{(l)}$, and likewise $\mathbf{c}_k^{(l)} := [c_k^{(l)} \; : \; \bar{\mathbf{c}}_k^{(l)T}]^T$. Eq. (4) can be re-written now as:

$$\bar{\mathbf{y}}_k^{(l)} = \mathbf{h}_k^{(l)} c_k^{(l)} + \bar{\mathbf{H}}_k^{(l)} \bar{\mathbf{c}}_k^{(l)} + \bar{\mathbf{w}}_k^{(l)}. \qquad (5)$$

Because $\bar{\mathbf{H}}_k^{(l)}$ has dimension $N \times (M-K-l+1)$, a left null vector exists for all l, when $N \geq M-K+1$. In the system with less receive- than transmit-antennas ($N < M$), the information block size K can be adjusted to satisfy $N \geq M-K+1$, so that the nulling/cancelling algorithm can always remain operational. Hence, LST-CR allows for any number of transmit- and receive-antennas. With $N \geq M-K+1$, left-multiplying both sides of (5) by the first row of the pseudoinverse $\mathbf{H}_k^{(l)\dagger}$, call it $\bar{\boldsymbol{\nu}}_k^{(l)T}$, we obtain

$$\bar{y}_k^{(l)} := \bar{\boldsymbol{\nu}}_k^{(l)T} \bar{\mathbf{y}}_k^{(l)} = \bar{\boldsymbol{\nu}}_k^{(l)T} \mathbf{h}_k^{(l)} c_k^{(l)} + \bar{\boldsymbol{\nu}}_k^{(l)T} \bar{\mathbf{w}}_k^{(l)}, \qquad (6)$$

where $\bar{\boldsymbol{\nu}}_k^{(l)T}$ satisfies $\bar{\boldsymbol{\nu}}_k^{(l)T}\mathbf{h}_k^{(l)} = 1$, and $\bar{\boldsymbol{\nu}}_k^{(l)T}\bar{\mathbf{H}}_k^{(l)} = \mathbf{0}^T$. To facilitate our performance analysis in Section V, we normalize $\bar{y}_k^{(l)}$ in (6) by $\|\bar{\boldsymbol{\nu}}_k^{(l)T}\|$, and re-write (6) as:

$$y_k^{(l)} := \frac{\bar{y}_k^{(l)}}{\|\bar{\boldsymbol{\nu}}_k^{(l)T}\|} := X_k^{(l)} c_k^{(l)} + \boldsymbol{\nu}_k^{(l)T} \bar{\mathbf{w}}_k^{(l)}, \qquad (7)$$

where $\boldsymbol{\nu}_k^{(l)T} := \bar{\boldsymbol{\nu}}_k^{(l)T}/\|\bar{\boldsymbol{\nu}}_k^{(l)T}\|$, and $X_k^{(l)}$ is defined as:

$$X_k^{(l)} := \frac{\bar{\boldsymbol{\nu}}_k^{(l)T}}{\|\bar{\boldsymbol{\nu}}_k^{(l)T}\|}\mathbf{h}_k^{(l)} = \frac{1}{\|\bar{\boldsymbol{\nu}}_k^{(l)T}\|}. \quad (8)$$

Repeating the nulling step (7) for the snapshots $k \in [1, K]$, and concatenating the resulting $y_k^{(l)}$ into $\mathbf{y}^{(l)} := [y_1^{(l)}, y_2^{(l)}, \ldots, y_K^{(l)}]^T$, we find:

$$\mathbf{y}^{(l)} = \mathbf{D}^{(l)}\boldsymbol{\Theta}\mathbf{s}^{(l)} + \mathbf{w}^{(l)}, \quad (9)$$

where $\mathbf{D}^{(l)} := \text{diag}[\boldsymbol{\nu}_1^{(l)T}\mathbf{h}_1^{(l)}, \ldots, \boldsymbol{\nu}_K^{(l)T}\mathbf{h}_K^{(l)}]$, and $\mathbf{w}^{(l)} := [\boldsymbol{\nu}_1^{(l)}\bar{\mathbf{w}}_1^{(l)T}, \boldsymbol{\nu}_2^{(l)}\bar{\mathbf{w}}_2^{(l)T}, \ldots, \boldsymbol{\nu}_K^{(l)}\bar{\mathbf{w}}_K^{(l)T}]^T$.

(S2) *Decoding*: To decode $\mathbf{s}^{(l)}$ from $\mathbf{y}^{(l)}$ in (9), we adopt the SD algorithm [1]. The complexity of SD for detecting a real $K \times 1$ block $\mathbf{s}^{(l)}$ is about $O(K^\alpha)$, where $\alpha \approx 6$ [1]. But recently, it has been claimed that on the average, and over a wide range of SNR, $\alpha \approx 3$ [4].

(S3) *Cancelling*: After obtaining $\hat{\mathbf{s}}^{(l)}$ from (S2), in this step $\hat{\mathbf{s}}^{(l)}$ is cancelled from $\bar{\mathbf{y}}_k^{(l)}$ via: $\bar{\mathbf{y}}_k^{(l+1)} = \bar{\mathbf{y}}_k^{(l)} - \mathbf{h}_k^{(l)}\boldsymbol{\theta}_k^T\hat{\mathbf{s}}^{(l)}$.

(S4) Repeating (S1)-(S3) until all L layers $\mathbf{c}^{(l)}$ have been decoded, completes the algorithm.

If we apply SD to decode all L layers $\mathbf{c}^{(l)}$, the complexity of the sub-optimum decoding is approximately $O(M^4) + O(LK^\alpha)$. Increasing K will increase decoding complexity. Compared with V-BLAST, the present scheme incurs only slightly higher decoding complexity, since K is chosen to be small compared to M.

V. Performance analysis

In this section, we will analyze the performance of the optimum and sub-optimum decoding schemes, relying on pairwise error probability (PEP) arguments.

A. Optimum Decoding

To benchmark the performance of LST-CR, we present (without proof) our first result on the maximum achievable diversity order when the receiver relies on ML decoding.

Proposition 1 *(Maximum Diversity gains with optimum decoding) If $\boldsymbol{\Theta}$ is selected so that $\prod_{k=1}^K \boldsymbol{\theta}_k^T\mathbf{e}^{(l)} \neq 0$ for any $\mathbf{e}^{(l)} := \mathbf{s}^{(l)} - \tilde{\mathbf{s}}^{(l)} \neq \mathbf{0}$ with $l = 1, \ldots, L$, then LST-CR achieves diversity gain equal to KN.*

Having established the diversity gain based on ML decoding, it is of interest to study the performance of low-complexity sub-optimum decoding alternatives.

B. LST-CR Decoding without error propagation

To investigate the performance of the LST-CR sub-optimum decoder, we start with the simpler case which assumes that no error propagates from layer to layer. Without loss of generality, we take our decoding order to be from the first to the Lth layer. Suppose we are to decode the lth layer, assuming that the first $(l-1)$ layers have been detected perfectly.

We will rely on $\mathbf{y}^{(l)}$ in (9) to perform ML detection of $\mathbf{s}^{(l)}$. Similar to [9], we start with the pairwise layer error event $\{\mathbf{s}^{(l)} \to \tilde{\mathbf{s}}^{(l)}\}$ as the event that the receiver decodes $\tilde{\mathbf{s}}^{(l)}$ erroneously, when $\mathbf{s}^{(l)}$ is actually sent. Without error propagation, (A2) and $\|\boldsymbol{\nu}_k^{(l)T}\| = 1$ imply that $\mathbf{w}^{(l)} \sim \mathcal{CN}(0, N_0\mathbf{I}_M)$ for all l in (9). Applying the inequality $\mathcal{Q}(x) \leq \exp(-x^2/2)$, we can upper bound the conditional PEP by

$$P(\mathbf{s}^{(l)} \to \tilde{\mathbf{s}}^{(l)}|\mathbf{D}^{(l)}) \leq \exp\left(-\frac{\bar{\gamma}}{4}\sum_{k=1}^K |X_k^{(l)}|^2|\boldsymbol{\theta}_k^T\mathbf{e}^{(l)}|^2\right), \quad (10)$$

where $\bar{\gamma} := 1/N_0$ is the average SNR per symbol with average energy 1.

To average the PEP in (10), we need the distribution of $|X_k^{(l)}|^2$. The latter can be found after recalling that $\mathbf{H}_k^{(l)}$ is an $N \times (M - K - l + 2)$ random matrix whose entries are i.i.d. complex Gaussian with zero mean and unit variance. It turns out that $\{|X_k^{(l)}|^2\}_{k=1}^K$ in (8) are chi-squared random variables with $2(N + K + l - M - 1)$ degrees of freedom [7]. The vector $\mathbf{x} := [|X_1^{(l)}|^2, \ldots, |X_k^{(l)}|^2]$ has mean $2(N+K+l-M-1)[1, \ldots, 1]$, and correlation matrix $\mathbf{R}^{(l)}$. Averaging both sides of (10) with respect to $\{|X_k^{(l)}|^2\}_{k=1}^K$, we obtain

$$P\left(\mathbf{s}^{(l)} \to \tilde{\mathbf{s}}^{(l)}\right) \leq E\{\exp(-\frac{\bar{\gamma}}{4}\sum_{k=1}^K |X_k^{(l)}|^2|\boldsymbol{\theta}_k^T\mathbf{e}^{(l)}|^2)\}. \quad (11)$$

Using [6, Eqn. (11)], we can write (11) more compactly as

$$P\left(\mathbf{s}^{(l)} \to \tilde{\mathbf{s}}^{(l)}\right) \leq \det\left(I + \frac{\bar{\gamma}}{4}\mathbf{D}_e\mathbf{M}\right)^{-(N+K+l-M-1)}, \quad (12)$$

where $\mathbf{D}_e := \text{diag}[|\boldsymbol{\theta}_1^T\mathbf{e}^{(l)}|^2, \ldots, |\boldsymbol{\theta}_K^T\mathbf{e}^{(l)}|^2]$, and $\mathbf{M}$ is a positive semidefinite matrix with $[\mathbf{M}]_{ij} = 2\sqrt{[\mathbf{R}^{(l)}]_{ij}}$.

If any two or more $X_k^{(l)}$'s are not fully correlated, then $\mathbf{M}$ has full rank [6]. We assume that $\mathbf{M}$ is full rank here, which we have also verified through simulations. Under this assumption, the eigenvalues of $\mathbf{M}$ are strictly positive since $\mathbf{M}$ is positive definite. Also $\mathbf{D}_e$ is a positive semidefinite matrix by definition; thus, the eigenvalues $\{\hat{\lambda}_k\}_{k=1}^K$ of $\mathbf{D}_e\mathbf{M}$ are all nonnegative.

Defining $\sigma := N + K - M - 1$, we obtain from (12)

$$P\left(\mathbf{s}^{(l)} \to \tilde{\mathbf{s}}^{(l)}\right) \leq \prod_{k=1}^K \left[1 + \frac{\bar{\gamma}}{4}\hat{\lambda}_k\right]^{-(\sigma+l)}. \quad (13)$$

Keeping in mind that $\{\hat{\lambda}_k\}_{k=1}^K$ depend on $\mathbf{e}^{(l)}$, we define the set $\mathcal{K}(\mathbf{e}^{(l)}) := \{k : \hat{\lambda}_k \neq 0\}$ with cardinality $|\mathcal{K}(\mathbf{e}^{(l)})|$. For sufficiently large values of $\bar{\gamma}$, we can approximate $P(\mathbf{s}^{(l)} \to \tilde{\mathbf{s}}^{(l)})$ in (12) as:

$$P\left(\mathbf{s}^{(l)} \to \tilde{\mathbf{s}}^{(l)}\right) \leq \left(\frac{\bar{\gamma}}{4}\right)^{-|\mathcal{K}(\mathbf{e}^{(l)})|(\sigma+l)}\left[\prod_{k\in\mathcal{K}(\mathbf{e}^{(l)})}\hat{\lambda}_k\right]^{-(\sigma+l)}. \quad (14)$$

From (14), we deduce that the diversity gain of the lth layer denoted by $G_{\rm d}^{(l)}$ is given by $G_{\rm d}^{(l)} = \mu(\sigma + l)$, where $\mu := \min_{\mathbf{e}^{(l)}\neq \mathbf{0}} |\mathcal{K}(\mathbf{e}^{(l)})|$ is the minimum value of $|\mathcal{K}(\mathbf{e}^{(l)})|$ over all distinct pairs $\{\mathbf{s}^{(l)}, \tilde{\mathbf{s}}^{(l)}\}$ [9]. For a given value of μ, we define the *product distance*, and the coding gain for layer l, respectively, as:

$$d(\mathbf{e}^{(l)}) := \left[\prod_{k\in\mathcal{K}(\mathbf{e}^{(l)})} \hat{\lambda}_k\right]^{\frac{1}{\mu}} \quad \text{and} \quad G_{\rm c}^{(l)} := \min_{\mathbf{e}^{(l)}\neq\mathbf{0}} d(\mathbf{e}^{(l)}).$$

Recalling the definition of $\mathcal{K}(\mathbf{e}^{(l)})$, we infer that the maximum $\mu = \max\{\min_{\mathbf{e}^{(l)}\neq\mathbf{0}} |\mathcal{K}(\mathbf{e}^{(l)})|\} = K$ is achieved. It can be easily seen that the necessary and sufficient condition for $\mu = K$ is that $\mathbf{D}_e$ has full rank for any $\mathbf{e}^{(l)} \neq \mathbf{0}$, i.e.,

$$\det(\mathbf{D}_e) = \prod_{k=1}^{K} \left|\boldsymbol{\theta}_k^T \mathbf{e}^{(l)}\right|^2 \neq 0, \quad \forall \mathbf{e}^{(l)} \neq \mathbf{0}. \qquad (15)$$

For $\mu = K$, the diversity gain of the lth layer is $G_{\rm d}^{(l)} = K(\sigma + l)$, while the coding gain of the lth layer is:

$$G_{\rm c}^{(l)} = \min_{\mathbf{e}^{(l)}\neq\mathbf{0}} \left[\prod_{k=1}^{K} \left|\boldsymbol{\theta}_k^T(\mathbf{s}^{(l)} - \tilde{\mathbf{s}}^{(l)})\right|^2 \prod_{k=1}^{K} \lambda_k\right]^{\frac{1}{K}}, \qquad (16)$$

where $\{\lambda_k\}_{k=1}^K$ are the eigenvalues of $\mathbf{M}$. Notice that $G_{\rm c}^{(l)}$ depends on the *minimum product distance* $\delta := \min_{\mathbf{e}^{(l)}\neq\mathbf{0}} \prod_{k=1}^K |\boldsymbol{\theta}_k^T(\mathbf{s}^{(l)} - \tilde{\mathbf{s}}^{(l)})|$. Although λ_k's affect system performance, they do not depend on the CR precoder $\boldsymbol{\Theta}$.

Let us now denote with $P_{\rm bl}^{(l)}$ the block error rate (BLER) of the lth layer. The $P_{\rm bl}^{(l)}$ can be upper-bounded by the union bound as $P_{\rm bl}^{(l)} \leq \sum_{\mathbf{s}^{(l)}} \sum_{d(\mathbf{e}^{(l)})\neq\mathbf{0}} P(\mathbf{s}^{(l)}) P(\mathbf{s}^{(l)} \to \tilde{\mathbf{s}}^{(l)})$, where $P(\mathbf{s}^{(l)})$ denotes the probability that $\mathbf{s}^{(l)}$ is transmitted. For equi-probable transmissions, $P(\mathbf{s}^{(l)}) = 1/|\mathcal{A}_s|^K$, $\forall \mathbf{s}^{(l)}$. At high SNR, we can thus approximate $P_{\rm bl}^{(l)}$ as

$$P_{\rm bl}^{(l)} \approx \frac{1}{|\mathcal{A}_s|^K} \sum_{d(\mathbf{e}^{(l)})=G_{\rm c}^{(l)}} \left[\left(\frac{\bar{\gamma}}{4}\right) G_{\rm c}^{(l)}\right]^{-G_{\rm d}^{(l)}} = O\left(\bar{\gamma}^{-G_{\rm d}^{(l)}}\right). \qquad (17)$$

Based on (17), we observe that $G_{\rm d}^{(l)}$ determines the slope of the BER-SNR curve. Hence, it is reasonable to maximize $G_{\rm d}^{(l)}$ first. When the diversity gains are fixed, the coding gains $\{G_{\rm c}^{(l)}\}_{l=1}^L$ measure the shift in SNR of the linearly precoded system as compared to the benchmark system with BER $= (\bar{\gamma}/4)^{-G_{\rm d}^{(l)}}$, at high SNR. Within the class of maximum diversity precoders, $\boldsymbol{\Theta}$ should be selected to maximize the coding gains $\{G_{\rm c}^{(l)}\}_{l=1}^L$ as well.

C. Asymptotic analysis with error propagation

In practice, error propagation is present and affects the overall system performance considerably. Although exact performance analysis is desirable, it is extremely difficult (if at all possible). Instead, we hereafter resort to an approximate asymptotic error analysis that accounts for error propagation.

With error propagation, $\mathbf{w}^{(l)}$ in (9) is a combination of the estimation error of the previously decoded layers and AWGN. Let $\hat{\mathbf{s}}^{(l)}$ be an estimate of $\mathbf{s}^{(l)}$ with estimation error $\tilde{\mathbf{e}}^{(l)} := \mathbf{s}^{(l)} - \hat{\mathbf{s}}^{(l)}$. Clearly, there is no error propagation issue when decoding the first layer. Focusing our attention to the l_0th layer's performance, we define the following error events for $l_0 > 1$: $\mathcal{G}_1 := \{\tilde{\mathbf{e}}^{(1)} \neq \mathbf{0}\}$, $\mathcal{G}_l := \{\tilde{\mathbf{e}}^{(1)} = \mathbf{0}, \cdots, \tilde{\mathbf{e}}^{(l_0-l)} = \mathbf{0}, \tilde{\mathbf{e}}^{(l_0+1-l)} \neq \mathbf{0}\}$, $l \in (1, l_0 - 1]$, and $\mathcal{G}_{l_0} := \{\tilde{\mathbf{e}}^{(1)} = \mathbf{0}, \cdots, \tilde{\mathbf{e}}^{(l_0-1)} = \mathbf{0}\}$.

According to the definition of $\{\mathcal{G}_l\}_{l=1}^{l_0}$, the BLER $P_{\rm bl}^{(l)}$ of the lth layer without error propagation is nothing but $P(\mathcal{G}_l)$. Let $\hat{P}_{\rm bl}^{(l_0)}$ be the BLER of the l_0 layer with error propagation. Defining $\pi_l := P(\tilde{\mathbf{e}}^{(l_0)} \neq \mathbf{0}|\mathcal{G}_l)$, we can express $\hat{P}_{\rm bl}^{(l_0)}$ as $\hat{P}_{\rm bl}^{(l_0)} := P(\tilde{\mathbf{e}}^{(l_0)} \neq \mathbf{0}) = \sum_{l=1}^{l_0} \pi_l P(\mathcal{G}_l)$. Also by (17), we have $P_{\rm bl}^{(l)} = O\left(\bar{\gamma}^{-\mu(N+K+l-M-1)}\right)$. Recalling the fact that $P_{\rm bl}^{(l)} = P(\mathcal{G}_l)$, we have the following approximation for sufficiently large $\bar{\gamma}$'s: $\hat{P}_{\rm bl}^{(l_0)} = O\left(\bar{\gamma}^{-\mu(N+K-M)}\right)$. Furthermore, the overall BLER can be written as:

$$\bar{P}_{\rm bl} = \frac{1}{M-K+1} \sum_{l=1}^{L} \hat{P}_{\rm bl}^{(l)} = O\left(\bar{\gamma}^{-\mu(N+K-M)}\right).$$

The diversity gain of the overall system is roughly the same as that of the first decoding layer, which is $G_{\rm d}^{(1)} = N - M + K$ in our case. This asymptotic analysis reveals that improving the performance of a few initially decoded layers is key to improving the overall system performance.

C.1 Decoding Order

So far we assumed that our decoding order is from the first to the Lth layer. Since BER is a monotone decreasing function of SNR, the decoding order should follow the order of post-detection SNR (see also [3]). Supposing perfect cancellation of the previously detected layers, i.e., $\mathbf{e}^{(l')} = \mathbf{0}$ for $l' = 1, \ldots, l-1$, we can treat $\mathbf{w}^{(l)}$ as AWGN in (9). Defining $\rho^{(l)}$ as the post-detection SNR of the lth layer, we have from (9)

$$\rho^{(l)} := \frac{E\{\mathrm{Tr}(\mathbf{D}^{(l)}\boldsymbol{\Theta}\mathbf{s}^{(l)}\mathbf{s}^{(l)\mathcal{H}}\boldsymbol{\Theta}^{\mathcal{H}}\mathbf{D}^{(l)\mathcal{H}})\}}{E\{\mathrm{Tr}(\mathbf{w}^{(l)}\mathbf{w}^{(l)\mathcal{H}})\}}. \qquad (18)$$

We notice that without error propagation, the denominator of $\rho^{(l)}$ is a constant. When $\boldsymbol{\Theta}$ is unitary, the numerator of (18) can be expressed as

$$E\{\mathrm{Tr}(\mathbf{D}^{(l)}\boldsymbol{\Theta}\mathbf{s}^{(l)}\mathbf{s}^{(l)\mathcal{H}}\boldsymbol{\Theta}^{\mathcal{H}}\mathbf{D}^{(l)\mathcal{H}})\} = \sum_{k=1}^{K} \frac{1}{\|\bar{\boldsymbol{\nu}}_k^{(l)T}\|^2}. \qquad (19)$$

Eq. (19) suggests that one should first decode the layer with the largest value of the sum $\sum_{k=1}^K 1/\|\bar{\boldsymbol{\nu}}_k^{(l)T}\|^2$.

VI. Design of Linear Precoders

In this section, we will design linear precoders that maximize diversity and coding gains of the sub-optimum layered decoding algorithm. To achieve maximum diversity and coding gains, we will follow the criteria suggested by the performance analysis of Section IV-B.

(CR1) *Diversity gain criterion:* Select $\mathbf{\Theta}$ such that: $\prod_{k=1}^{K}|\boldsymbol{\theta}_k^T(\mathbf{s}^{(l)}-\tilde{\mathbf{s}}^{(l)})|\neq 0,\ \forall \mathbf{s}^{(l)}\neq\tilde{\mathbf{s}}^{(l)},\quad \forall l\in[1,M-K+1]$. When $\mathbf{\Theta}$ is chosen to satisfy (15), we infer that $\mu=K$, and hence $G_d^{(l)}=K(N+K+l-M-1)$.

(CR2) *Coding gain criterion:* Select $\mathbf{\Theta}$ to maximize the product distance $\delta:=\min_{\mathbf{s}^{(l)}\neq\tilde{\mathbf{s}}^{(l)}}\prod_{k=1}^{K}|\boldsymbol{\theta}_k^T(\mathbf{s}^{(l)}-\tilde{\mathbf{s}}^{(l)})|$.

With $\mathbf{\Theta}$ chosen to satisfy (CR1) and (CR2), constellation-rotating codes have been designed in [10,11], without the LST structure. The novelty here consists of designing the CR linear precoders not only for maximum diversity gains, but also for maximum coding gains *per layer*. In general, there are two design methodologies for constructing CR precoders [2,5,11].

A. Parameterization of unitary precoders

So long as the constellation set $\mathcal{A}_s$ is finite, there exists at least one unitary precoder $\mathbf{\Theta}$ ensuring maximum diversity gain [10]. Because unitary CR matrices accept a finite parameterization, we can narrow our search from the class of linear precoders to the unitary ones.

Proposition 2 *(Parameterization of unitary matrices) The $K\times K$ unitary precoder $\mathbf{\Theta}$ can be constructed as: $\mathbf{\Theta}=\prod_{1\leq k\leq K-1,k+1\leq l\leq K}\mathbf{G}_{kl}(\theta_{kl},\phi_{kl})$, where $\mathbf{G}_{k,l}(\theta_{kl},\phi_{kl})$ is a complex Givens matrix , which is similar to the identity matrix with (k,k) and (l,l) entries replaced by $\cos\theta_{kl}$, and (k,l) and (l,k) entries replaced by $\sin\theta_{kl}e^{-i\phi_{kl}}$ and $-\sin\theta_{kl}e^{-i\phi_{kl}}$, respectively, with $\theta_{kl}\in[-\pi,\pi]$ and $\phi_{kl}\in[-\pi/2,\pi/2]$.*

In selecting a unitary $\mathbf{\Theta}$ to maximize diversity and coding gains, Proposition 2 implies that there are $K(K-1)$ parameters to be optimized. For small values of K (e.g., $K\leq 3$) and small constellation sizes (BPSK or 4-QAM), an exhaustive search is feasible. However, when the constellation size is large or when K increases, the search space becomes prohibitively large. Then, the algebraic approach summarized next is more desirable from a design point of view.

B. Algebraic construction of CR precoders

The square CR matrices designed in this subsection are allowed to have any size K. Earlier constructions in [5] follow as special cases of our design with $K=2^Q$, and $Q\in\mathbb{Z}$. The general form of our CR matrices has row-wise Vandermonde structure [10,11]

$$\mathbf{\Theta}=\frac{1}{\psi}\begin{bmatrix}1 & \alpha_0 & \cdots & \alpha_0^{K-1}\\ 1 & \alpha_1 & \cdots & \alpha_1^{K-1}\\ \vdots & \vdots & & \vdots\\ 1 & \alpha_{K-1} & \cdots & \alpha_{K-1}^{K-1}\end{bmatrix}, \qquad (20)$$

where ψ is a normalizing scalar that enforces the power constraint: $E(\|\mathbf{\Theta s}\|^2)=E(\|\mathbf{s}\|^2)=K$. The distinct generators $\{\alpha_k\}_{k=0}^{K-1}$ are chosen depending on K, as follows:

(C1) If K is an Euler number[1]; i.e., $K\in\mathcal{S}_1=\{\phi(P):P\not\equiv 0\ (\mathrm{mod}\ 4)\}$, then $\{\alpha_k\}_{k=1}^{K}$ are chosen to be roots of the equation $\Phi_P(x)=0$, where $\Phi_P(x)=\prod_{i\in\mathcal{I}}(x-\alpha^i)$ with $\mathcal{I}:=\{i\mid \gcd(i,P)=1\ \text{and}\ i\in[1,P)\}$ and $\alpha:=e^{j2\pi/P}$.

(C2) If K is a power of 2; i.e., $K\in\mathcal{S}_2:=\{2^Q:Q\in\mathbb{N}\}$, then $\{\alpha_k\}_{k=1}^{K}$ are chosen as the roots of $x^K-j=0$. The precoder (20) then becomes a unitary precoder that can be factored as: $\mathbf{\Theta}=\mathbf{F}_K^T\mathrm{diag}\left(1,\alpha,\ldots,\alpha^{K-1}\right)$, where $\mathbf{F}_K$ is the K-point inverse fast Fourier transform (IFFT) matrix with $(m+1,n+1)$st entry $K^{-1/2}\exp(j2\pi mn/K)$, and $\alpha:=\exp(j\pi/(2K))$.

(C3) If $K\notin\mathcal{S}_1\cup\mathcal{S}_2$, then $\{\alpha_k\}_{k=1}^{K}$ are chosen to be the roots of $x^K-(1+j)=0$.

It can be shown that when $K\in\{\mathcal{S}_1\cup\mathcal{S}_2\}$, the CR precoder $\mathbf{\Theta}$ in (20) achieves maximum diversity and coding gains over any QAM constellation [11].

VII. Design Examples and Simulations

In this section, we present simulated performance results using QPSK constellations.

Test Case 1: (LST-CR vs. V-BLAST) We consider a multi-antenna system with $M=N=15$, $K=2$, and simulate LST-CR at transmission rate 28 bps/Hz. Compared with the uncoded V-BLAST, the transmission efficiency here η is $14/15$. In LST-CR, we only decode the first four (out of 14) layers with SD to further reduce the decoding complexity. To maintain the same transmission rate with V-BLAST for $M=N=15$, we also simulate LST-CR at transmission rate 30 bps/Hz by adopting 16-QAM constellations at the last decoding layer. Compared with the V-BLAST, LST-CR in Fig. 2 shows about 12 dB gain at BER $=10^{-3}$. Also we observe that the BER-SNR curves of LST-CR have different slope from those of V-BLAST with $M=14, N=15$, which confirms that CR precoding per layer enhances the transmit diversity gains.

Test Case 2: (LST-CR vs. LD) Fig. 3 depicts performance comparisons between LST-CR and LD codes with $M=8$ and $N=4$. We apply the nulling/cancelling algorithm to decode LD, and apply sub-optimum decoding that relies on nulling/cancelling and SD to decode LST-CR. At low to moderate SNR ($5-15$ dB), LST-CR performs similar to LD codes, but at moderate to high SNR ($15-30$ dB), LST-CR codes outperform their LD counterparts considerably. At BER $=10^{-5}$, LST-CR gains about $4-5$ dB over LD codes.

Test Case 3: (Effects of decoding order) In this simulation, we test how the layer ordering can affect the performance of a system with $M=N=15$. Fig. 4 shows that at BER $=10^{-4}$, the layer ordering can improve performance by $3-4$ dB, which verifies our claims in subsection V-C.1.

Test Case 4: (Effects of precoder size) In a system with $M=N=15$, we compare BER performance for various values of K in Fig. 5. It can be easily seen that the diversity increases as K increases.

[1] An Euler number $\phi(P)$ is defined as the number of positive integers $<P$ which are relatively prime to P.

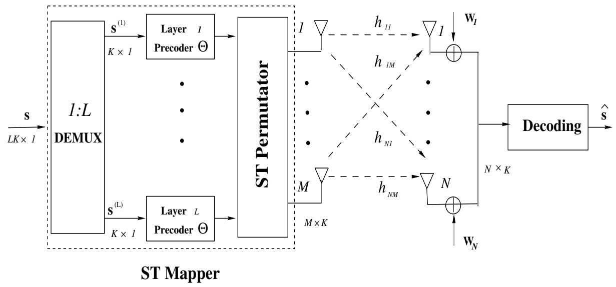

Fig. 1. LST-CR Model

VIII. Conclusion

A novel linearly precoded layered space-time scheme was derived to achieve reliable high rate transmissions for any number of transmit- and receive- antennas. It was shown flexible to trade off performance with rate and decoding complexity. A sub-optimum decoder was proposed with complexity slightly higher than the nulling/cancelling in V-BLAST scheme. Decoding order was utilized to improve performance. A systematic design of linear constellation-rotating precoders was exploited to achieve maximum coding gains per layer. Simulation results corroborated our theoretical analyses.

Acknowledgment: The authors would like to thank Drs. B. Hassibi and B. Hochwald for providing their LD codes.

References

[1] M. O. Damen, A. Chkeif, and J.-C. Belfiore, "Lattice codes decoder for space-time codes," *IEEE Communications Letters*, vol. 4, no. 5, pp. 161–163, May 2000.

[2] V. M. DaSilva and E. S. Sousa, "Fading-resistant modulation using several transmitter antennas," *IEEE Transactions on Communications*, vol.45, no. 10, pp. 1236-1244, Oct. 1997.

[3] 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," *IEEE JSAC*, vol. 17, no. 11, pp. 1841-1852, Nov. 1999.

[4] B. Hassibi and B. Hochwald, "High-rate codes that are linear in space and time," *IEEE Trans. on Info. Theory*, Aug. 2000 (submitted); downloadable from `http://mars.bell-labs.com`

[5] X. Giraud, E. Boutillon, and J.-C. Belfiore, "Algebraic tools to build modulation schemes for fading channels," *IEEE Trans. on Info. Theory*, vol. 43, no. 3, pp. 938-952, May 1997.

[6] P. Lombardo, G. Fedele, and M. M. Rao, "MRC performance for binary signals in Nakagami fading with general branch correlation," *IEEE Trans. on Communications*, vol. 47, no. 1, pp. 44-52, Jan. 1999.

[7] D. Shiu and J. M. Kahn, "Design of high-throughput codes for multiple-antenna wireless systems," *IEEE Trans. on Information Theory*, Jan. 1999 (submitted); downloadable from `http://www.cs.berkeley.edu/~jmk//res.areas/`

[8] V. Tarokh, H. Jafarkhani, and A. R. Calderbank, "Space-time block codes from orthogonal designs," *IEEE Trans. on Info. Theory*, vol. 45, no. 5, pp. 1456-1466, July 1999.

[9] V. Tarokh, N. Seshadri, and A. Calderbank, "Space-time codes for high data rate wireless communications: performance criterion and code construction," *IEEE Trans. on Info. Theory*, vol. 44, no. 2, pp. 744–765, Mar. 1998.

[10] Y. Xin, Z. Wang, and G. B. Giannakis, "Space-time diversity systems based on linear constellation precoding," *IEEE JSAC*, Mar. 2001 (submitted); downloadable from `http://spincom.ece.umn.edu/zhiqiang/xiwg01j.ps`

[11] Y. Xin, Z. Wang, and G. B. Giannakis, "Space-time constellation-rotating codes maximizing diversity and coding gains," in *Proc. of GLOBECOM*, San Antonio, TX, Nov. 25-29, 2001, pp. 455-459.

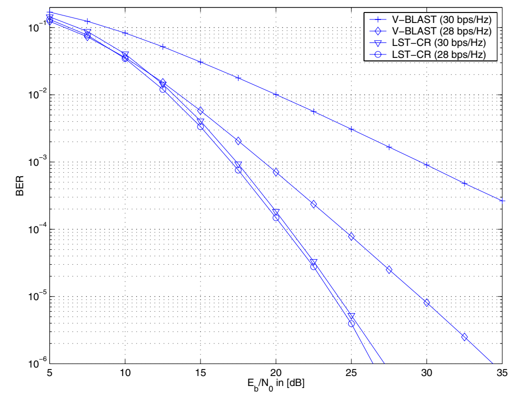

Fig. 2. LST-CR versus V-BLAST $(M = N = 15, K = 2)$

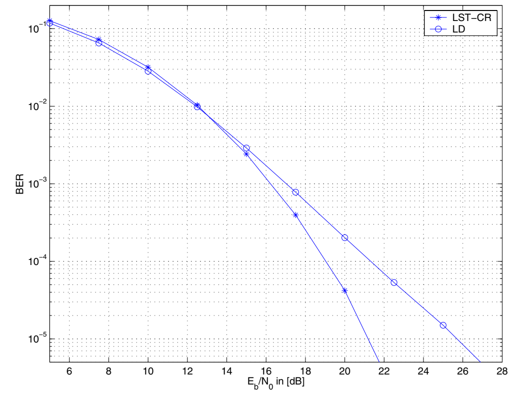

Fig. 3. LST-CR versus LD $(M = 8, N = 4)$

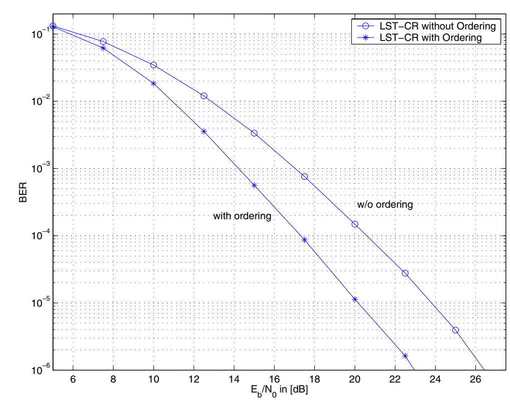

Fig. 4. Decoding of LST-CR with and without ordering

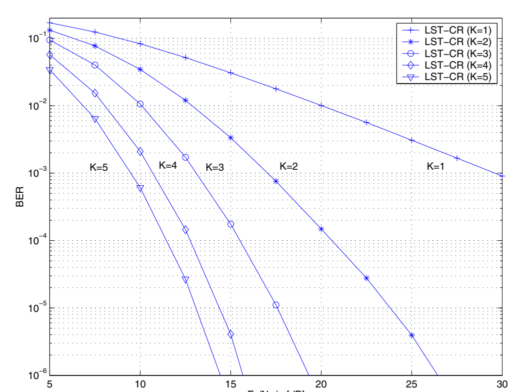

Fig. 5. LST-CR with variable redundancy and diversity gain