

Turbo Space-Time Codes with Time Varying Linear Transformations

Hangjun Chen and Alexander Haimovich

Department of Electrical and Computer Engineering

New Jersey Institute of Technology

Newark, NJ 07102

Email: {hangjun.chen; alexander.m.haimovich}@njit.edu

Abstract

Turbo space-time codes with symbols precoded by randomly chosen unitary *time variant linear transformations* (TVLT) are investigated in this paper. It is shown that turbo codes with TVLT achieve full diversity gain and do not require exhaustive tests of the rank criterion. We prove that the coding performance of turbo space-time codes with TVLT improves with the Hamming distance between codewords (number of different columns). As an additional benefit of the application of TVLT, with the removal of the constant modulation condition, we prove that throughput rates achieved by these codes are significantly higher than those for conventional space-time codes. Finally, we develop an EXIT chart analysis for turbo space-time codes with TVLT. The EXIT chart is used to facilitate the design of the codes. It is noted that so far, turbo space-time codes have resisted EXIT chart analysis due to a large number of parameters needed to specify channel conditions. It is shown that with TVLT, the singular eigenvalues of the channel matrix are sufficient to predict the convergence region of the iterative decoder. Subsequently, this analysis also enables reasonably accurate estimates of the frame error rate (FER). Numerical results are provided to demonstrate diversity gains, coding gains and rates of turbo space-time codes with TVLT. The EXIT chart is explored for determining convergence and predicting FER performance.

I. INTRODUCTION

The invention of space-time codes [1] has been seminal in the advance of multiple input multiple output (MIMO) techniques. Space-time codes apply to multiple transmit antennas to

achieve both diversity and coding gains over fading channels. However, the design of full diversity codes with good coding gain has been challenging. One of the difficulties has been the lack of effective design tools to meet specified diversity and coding gains criteria. In most cases, to ensure meeting diversity and gain specifications [1], exhaustive checks are required of all pairs of codewords, and numerical simulations are needed to evaluate overall performance.

Turbo space-time codes (STC) are specially designed turbo codes for MIMO channels that can achieve much higher coding gains than conventional space-time trellis codes. Existing schemes such as [2], [3], and [4], etc., display similar performance but have distinct encoder/decoder structures. For all turbo-STC, exhaustive tests are required to ensure full diversity. If the test fails, the codes have to be adjusted, and the testing procedure is restarted.

The main goal in this paper is to simplify the design of full diversity turbo space-time codes. Phase sweeping [5], [6], [7], a precursor to space-time codes, guarantees full diversity, but no coding gain. It still attracts attention both in academia and industry [8], [9]. The approach taken in this paper is a technique related to phase sweeping and is referred to as *time varying linear transformation* (TVLT) [10]. TVLT offers more flexibility for integration in space-time codes than phase sweeping. With TVLT, transmitted symbols are modified by a unitary transformation represented by a unitary matrix. The elements of the matrix are in part selected at random and in part designed to meet the unitary conditions. The transformation changes at each time interval, but the sequence of transformations is known to both the transmitter and the receiver. A previous attempt of applying TVLT to turbo space-time codes proved unsuccessful.

Not only does TVLT guarantee full diversity, but it also provides coding gain with high probability. In particular, we will show that the probability of not meeting the coding gain specification vanishes exponentially with the Hamming distance between codewords defined as the number of different columns in pairs of codewords. Since turbo codes typically feature considerable Hamming distances, an exhaustive check of all pairs of codewords is not necessary.

Besides simplifying design, TVLT can also be used to construct full diversity space-time codes with higher rate than codes with constant modulation, e.g., 4-PSK, at each time interval. The maximum data rate limit derived in [1] for a full diversity space-time codes is based on the assumption of constant modulation. TVLT has the effect of varying the modulation at each time interval. It leads to a higher maximum rate that will be derived in the paper. For example, full diversity space-time codes with 8-PSK modulation have a maximum rate of 3 b/s/Hz. We will

demonstrate a full diversity TVLT 8-PSK code with a rate of 4 b/s/Hz.

Finally, we show that turbo STC with TVLT lend themselves to the application of a form of the *extrinsic information transfer* (EXIT) chart especially modified for the task. This method is an extension of the EXIT chart for single antenna turbo codes [11], [12], [13] to turbo STC. It provides an efficient, low complexity method to quickly predict the performance of turbo STC. With this convenient tool, we can easily compare between different designs and different codes and select the best one. Several new turbo STC are developed and analyzed in this manner.

The remaining of the paper is organized as follows. In Section II, after presenting the channel model, TVLT codes are introduced. Salient properties of turbo STC with TVLT are explored in Section III. The performance analysis based on the EXIT charts is developed in Section IV. Numerical results are presented in Section V and concluding remarks are found in Section VI.

II. CHANNEL MODEL AND TVLT

In this paper we consider MIMO systems with M transmit and N receive antennas over block Rayleigh fading channel. The received signal at time interval t is

$$\mathbf{r}(t) = \mathbf{H}\mathbf{x}(t) + \mathbf{n}(t), \quad (1)$$

where $\mathbf{x}(t)$ is the $M \times 1$ vector of the transmitted signal; $\mathbf{r}(t)$ is the $N \times 1$ received signal vector; \mathbf{H} is the $N \times M$ channel gain matrix, which consists of complex-valued scalars h_{ij} , $i = 1, \dots, N$, $j = 1, \dots, M$, modeled as zero-mean, mutually independent, identically distributed Gaussian random variables with unity variance. The term $\mathbf{n}(t)$ is zero-mean, additive white Gaussian noise (AWGN) with variance $M/(2 \text{SNR})$ per dimension, where SNR is the signal to noise ratio per symbol. With this model, the mean SNR at the receiver is independent of the number of transmit antennas.

The transmitter with TVLT is shown in Fig. 1. Denote a codeword produced by the encoder $\mathbf{C} = [\mathbf{c}(1), \dots, \mathbf{c}(L)]$, with L the codeword length. At each time interval t , $t = 1, \dots, L$, a vector of coded symbols $\mathbf{c}(t) = [c_1(t), \dots, c_M(t)]^T$ is multiplied by a predetermined $M \times M$ unitary matrix $\mathbf{G}(t)$ and yields that vector of symbols $\mathbf{x}(t)$, which is subsequently transmitted,

$$\mathbf{x}(t) = \mathbf{G}(t)\mathbf{c}(t). \quad (2)$$

The set of unitary matrices $\{\mathbf{G}(t)\}_{t=1}^L$ is generated randomly, independent of the codewords it is applied to. Once generated, the set is stored both at the transmitter and at the receiver

and it stays fixed for the duration of the communication. Insofar as the elements of $\mathbf{G}(t)$ are time-varying and realizations of random variables, $\mathbf{x}(t)$ appears to be subject to a time-varying, random modulation. At the receiver, a random modulation is considerably simpler to demodulate than decoding a random code since demapping is carried out symbol by symbol rather than on long sequences.

Next, we discuss the generation of these unitary matrices for an arbitrary number of transmit antennas. For $p, q = 1, \dots, M$, let $g_{p,q}(t)$ be the elements, and $\mathbf{g}_q(t)$ be the columns of $\mathbf{G}(t)$. Since $\mathbf{G}(t)$ is unitary, it must satisfy

$$\mathbf{g}_p^\dagger(t)\mathbf{g}_q(t) = \begin{cases} 1 & p = q \\ 0 & p \neq q \end{cases}, \quad (3)$$

where the superscript denotes transpose conjugate. Since the elements $g_{p,q}(t)$ are generally complex-valued, then each $\mathbf{G}(t)$ is specified by $2M^2$ real-valued parameters of which M^2 are prescribed by (3) and the rest are free, but constrained $|g_{p,q}(t)|^2 \leq \|\mathbf{g}_q(t)\|^2 = 1$. These requirements can be met by elements of the following form:

$$g_{p,q}(t) = \cos \alpha_{p,q}(t) e^{j\beta_{p,q}(t)},$$

where $\alpha_{p,q}(t), \beta_{p,q}(t) \in [0, 2\pi)$. Of these, M^2 are independent realizations picked from the uniform distribution over $[0, 2\pi)$ and the other M^2 terms are found from (3). Of particular interest in this paper is the case $M = 2$, for which we pick $\alpha_{11}(t), \beta_{11}(t), \beta_{12}(t), \beta_{21}(t)$ and solve (3) for the rest to obtain

$$\mathbf{G}(t) = \begin{bmatrix} \cos \alpha_{11}(t) e^{j\beta_{11}(t)} & \sin \alpha_{11}(t) e^{j\beta_{12}(t)} \\ -\sin \alpha_{11}(t) e^{j\beta_{21}(t)} & \cos \alpha_{11}(t) e^{j(-\beta_{11}(t) + \beta_{12}(t) + \beta_{21}(t))} \end{bmatrix}. \quad (4)$$

III. TURBO-STC WITH TVLT

In this section, properties of turbo space-time codes with TVLT are investigated. It is shown that full diversity can be practically assured. In addition, if the minimum Hamming distance between all pairs of space-time codewords is sufficiently large, the coding gain exceeds a prescribed threshold with high probability. The Hamming distance between two codewords is defined as the number of columns that are different between the codewords. Finally, the throughput rate of these codes is investigated and it is shown that it exceeds the maximum rate for conventional space-time codes derived in [1].

A. Diversity and Coding Gains

Let two distinct codewords of a space-time code be represented by the $M \times L$ matrices $\mathbf{C} = [\mathbf{c}(1)\dots\mathbf{c}(L)]$ and $\mathbf{B} = [\mathbf{b}(1)\dots\mathbf{b}(L)]$, respectively. According to the rank criterion, the code achieves full transmit diversity if for every possible pair of codewords \mathbf{C} and \mathbf{B} , the error matrix $\mathbf{E} = [\mathbf{e}(1), \dots, \mathbf{e}(L)]$, where $\mathbf{e}(t) = \mathbf{c}(t) - \mathbf{b}(t)$, has rank M [1]. For special space-time codes, simple proofs of full diversity can be found [1], however for the general case, ascertaining full diversity is a cumbersome process that requires testing the rank of all pairs of codewords. When the test fails, the encoder is redesigned and the test starts over. For turbo space-time codes, there are no explicit guidelines for designing full diversity codes. According to [3], without the tedious check-adjust process, a turbo space-time code with a randomly picked interleaver usually fails to achieve full diversity.

For a full diversity space-time code, the coding gain achieved by the pair \mathbf{B} , \mathbf{C} is $\xi = (\lambda_1, \dots, \lambda_M)^{1/M} = (\det(\mathbf{E}\mathbf{E}^\dagger))^{1/M}$, where $\lambda_1, \dots, \lambda_M$ denote the M eigenvalues of $\mathbf{E}\mathbf{E}^\dagger$ [1]. The coding gain of the space-time code is the minimum of ξ sought over all pairs of codewords.

Consider now a space-time code with TVLT. Let two codewords be $\overline{\mathbf{B}} = [\mathbf{G}(1)\mathbf{b}(1), \dots, \mathbf{G}(L)\mathbf{b}(L)]$ and $\overline{\mathbf{C}} = [\mathbf{G}(1)\mathbf{c}(1), \dots, \mathbf{G}(L)\mathbf{c}(L)]$, respectively, and the corresponding error matrix $\overline{\mathbf{E}} = [\mathbf{G}(1)\mathbf{e}(1), \dots, \mathbf{G}(L)\mathbf{e}(L)]$. Let d_h be the Hamming distance between \mathbf{B} and \mathbf{C} , i.e., the number of nonzero columns in \mathbf{E} . It follows immediately that d_h is also the number of non-zero columns in $\overline{\mathbf{E}}$, except for the rare occasion when $\overline{\mathbf{e}}(t) = \mathbf{G}(t)\mathbf{e}(t) = \mathbf{0}$, but $\mathbf{e}(t) \neq \mathbf{0}$. Since the elements of the transformations $\{\mathbf{G}(t)\}_{t=1}^L$ are realizations of random variables, with high probability, for $d_h \geq M$, $\text{rank } \overline{\mathbf{E}} = M$, thus giving the code full diversity.

The linear transformations $\mathbf{G}(t)$ affect the coding gain in a manner specified by the following theorem.

Theorem 1: Let \mathbf{B} and \mathbf{C} be a pair of $M \times L$ codewords with Hamming distance d_h . Let d_{\max}^2 be the largest Euclidean norm of all columns of \mathbf{E} , i.e., $d_{\max}^2 = \max_{t=1, \dots, L} \{\|\mathbf{e}(t)\|^2\}$, and d_{\min}^2 be the smallest norm of all nonzero columns of \mathbf{E} , i.e., $d_{\min}^2 = \min_{t=1, \dots, L} \{\|\mathbf{e}(t)\|^2, \|\mathbf{e}(t)\|^2 \neq 0\}$. With $\overline{\mathbf{B}}$, $\overline{\mathbf{C}}$ the corresponding pair of codewords following TVLT, define the error matrices $\mathbf{E} = \mathbf{C} - \mathbf{B}$ and $\overline{\mathbf{E}} = \overline{\mathbf{C}} - \overline{\mathbf{B}}$. Assume that the application of TVLT results in a full diversity code. If $\xi = \left(\det(\overline{\mathbf{E}}\overline{\mathbf{E}}^\dagger)\right)^{1/M}$ is the coding gain achieved by $\overline{\mathbf{B}}$, $\overline{\mathbf{C}}$, then $\text{Pr}\{\xi < u\}$, the probability that the coding gain is lower than a given threshold u , where $0 < u < \left(d_{\max}^2 \cdot (d_{\min}^2)^{M-1}\right)^{1/M}$,

decreases exponentially with $\lfloor d_h / (M - 1) \rfloor$, where $\lfloor x \rfloor$ denotes the largest integer less than x .

Proof: Without loss of generality, rearrange \mathbf{E} such that the nonzero columns are listed first. In particular, let the first column be the one with the largest norm d_{\max}^2 . Then, following TVLT, we have

$$\begin{aligned}\bar{\mathbf{E}} &= [\mathbf{G}(1)\mathbf{e}(1), \dots, \mathbf{G}(d_h)\mathbf{e}(d_h), 0, \dots, 0] \\ &\triangleq [\bar{\mathbf{e}}(1), \dots, \bar{\mathbf{e}}(d_h), 0, \dots, 0].\end{aligned}$$

Divide the columns $\{\bar{\mathbf{e}}(2), \dots, \bar{\mathbf{e}}(d_h)\}$ into $U = \lfloor d_h / (M - 1) \rfloor$ disjoint partitions, each containing $M - 1$ columns, and concatenate each partition with the first column of $\bar{\mathbf{E}}$ to construct U square matrices $\mathbf{V}(\mu)$, $\mu = 1, \dots, U$. Each matrix $\mathbf{V}(\mu)$ is generated by a set of $(M - 1)$ transformations $\mathbf{G}(t)$. Since the first column is common to all $\mathbf{V}(\mu)$, its transformation can be absorbed in the MIMO channel. Recalling that each $\mathbf{G}(t)$ is obtained from M^2 realizations of a random variable, it follows that its domain is a hypercube with volume $(2\pi)^{M^2}$. Extending this reasoning to $\mathbf{V}(\mu)$, its domain is a hypercube with volume $\Psi_\mu = (2\pi)^{(M-1)M^2}$.

The determinant $\det(\mathbf{V}(\mu)^\dagger \mathbf{V}(\mu))$ is a continuous function of the angles forming the transformations $\mathbf{G}(t)$, since its value is determined by trigonometric and exponential functions of these angles. Moreover, $\det(\mathbf{V}(\mu)^\dagger \mathbf{V}(\mu))$ is bounded by $0 \leq \det(\mathbf{V}(\mu)^\dagger \mathbf{V}(\mu)) \leq \prod_{k=1}^M (\mathbf{V}(\mu)^\dagger \mathbf{V}(\mu))_{kk}$, where $(\mathbf{V}(\mu)^\dagger \mathbf{V}(\mu))_{kk}$ is an element on the diagonal of the matrix $\mathbf{V}(\mu)^\dagger \mathbf{V}(\mu)$. The left hand side of the inequality follows from the non-negative property of a determinant of a Hermitian matrix, and the right hand side is a consequence of Hadamard's inequality. Direct evaluation shows that for any μ , $\prod_{k=1}^M (\mathbf{V}(\mu)^\dagger \mathbf{V}(\mu))_{kk} = d_{\max}^2 \prod_{t=(\mu-1)(M-1)+1}^{\mu(M-1)} \|\mathbf{e}(t)\|^2$. As the TVLT transformations $\mathbf{G}(t)$ range through all possible values, $\det(\mathbf{V}(\mu)^\dagger \mathbf{V}(\mu))$ maps onto all values between the two bounds.

Consequently, for any u , with $0 < u^M < d_{\max}^2 (d_{\min}^2)^{M-1} \leq d_{\max}^2 \prod_{t=(\mu-1)(M-1)+1}^{\mu(M-1)} \|\mathbf{e}(t)\|^2$, there exist a non-empty Υ_μ region in Ψ_μ such that if the domain of $\mathbf{V}(\mu)$, $\mathbf{A}_\mu \in \Upsilon_\mu$, then $\det(\mathbf{V}(\mu)^\dagger \mathbf{V}(\mu)) \geq u^M$. Let $\bar{\Upsilon}_\mu$ denote the complement region of Υ_μ , that is if $\mathbf{A}_\mu \in \bar{\Upsilon}_\mu$, then $\det(\mathbf{V}(\mu)^\dagger \mathbf{V}(\mu)) < u^M$. Then, if the volume of $\bar{\Upsilon}_\mu$ is denoted $|\bar{\Upsilon}_\mu|$,

$$\begin{aligned}\Pr \{ \det(\mathbf{V}(\mu)^\dagger \mathbf{V}(\mu)) < u^M \} \\ = \frac{|\bar{\Upsilon}_\mu|}{(2\pi)^{(M-1)M^2}} < 1.\end{aligned}\tag{5}$$

It is proved in the appendix that $\det(\overline{\mathbf{E}\mathbf{E}^\dagger}) \geq \det(\mathbf{V}(\mu)^\dagger \mathbf{V}(\mu))$, $\mu = 1, \dots, U$. It follows that,

$$\begin{aligned}
& \Pr\{\xi < u\} = \Pr\left\{\det(\overline{\mathbf{E}\mathbf{E}^\dagger}) < u^M\right\} \\
& \leq \Pr\left\{\det(\mathbf{V}(1)^\dagger \mathbf{V}(1)) < u^M, \dots, \det(\mathbf{V}(U)^\dagger \mathbf{V}(U)) < u^M\right\} \\
& = \prod_{\mu=1}^U \Pr\left\{\det(\mathbf{V}(\mu)^\dagger \mathbf{V}(\mu)) < u^M\right\} \\
& = \prod_{\mu=1}^U \frac{|\overline{\Upsilon}_\mu|}{(2\pi)^{(M-1)M^2}} \\
& \leq r^{\lfloor d_h/(M-1) \rfloor},
\end{aligned} \tag{6}$$

where

$$r = \max_{\mu=1, \dots, U} \left\{ \frac{|\overline{\Upsilon}_\mu|}{(2\pi)^{(M-1)M^2}} \right\}. \tag{7}$$

In (6), line 3 follows since $\{\mathbf{V}(\mu), \mu = 1, \dots, U\}$ are independent; in line 5, we recall that $U = \lfloor d_h/(M-1) \rfloor$. ■

Theorem 1 answers the question: what is the probability that the coding gain falls below a set value. It states that this probability vanishes exponentially with the ratio $\lfloor d_h/(M-1) \rfloor$. It follows that a desired feature of space-time codes with TVLT is a large Hamming distance.

Conventional space-time trellis codes, such as proposed in [1], generally do not satisfy a large Hamming distance requirement. That means that application of TVLT to such codes could potentially result in low coding gain. Hence, TVLT is not suitable to be used with conventional space-time trellis codes.

Although not much is known on efficient methods to calculate the minimum Hamming distances for arbitrary, non-binary turbo space-time codes, such codes usually feature relatively large minimum Hamming distances. Even when there are low weight codewords (leading to low Hamming distances), their number is small and they do not have a significant impact on performance [14]. The calculation of $d_{h,\min}$ for binary turbo codes is better known. For example, in [15] it is shown that for a rate 1/2, eight-state binary turbo code with convolutional component codes and interleaver size 1000, the upper bound on the minimum Hamming distance is about 100. Techniques have been published in literature for optimizing the Hamming distance in binary codes [16], [17], but it is unknown whether these methods extend to non-binary codes.

B. Rate

According to Theorem 3.3.1 and Corollary 3.3.1 in [1], for a M transmit and N receive antennas, and signal constellation size 2^b , the full diversity of MN can be achieved with a maximum rate of b bits/s/Hz. Inspection of the proof in the reference, reveals that the rate limit is a consequence of the restriction of constant modulation during the codeword. With TVLT, this restriction is removed, since the modulation constellation changes each time interval. Consequently, a new rate derivation is required for space-time codes with TVLT.

The next theorem completes the characterization of space-time codes with TVLT, by establishing the code rate and its relation to the Hamming distance.

Theorem 2: Consider a space-time code for M transmit antennas. Assume a signal constellation \mathcal{Q} (before TVLT application) with cardinality $|\mathcal{Q}| = 2^b$ elements. Let L denote the codeword length. Then the maximum rate of a code with minimum Hamming distance $d_{h,\min}$ is

$$R \leq Mb - \frac{1}{L} \log_2 V,$$

in bits per channel use, where

$$V = \sum_{j=0}^{d_{h,\min}} \binom{L}{j} (2^{Mb} - 1)^j.$$

Proof: Regard the columns of the $M \times L$ space-time codeword matrix as supersymbols defined over a superalphabet constellation \mathcal{Q}^M with $|\mathcal{Q}^M| = 2^{Mb}$. Each codeword can be viewed as an L -tuple in a L -dimensional space $[\mathcal{Q}^M]^L$. The sphere-packing bound [18] can be applied to obtain the maximum rate. Imagine a ball of radius $d_{h,\min}$ in the L -dimensional space, centered at a codeword point, including all points at Hamming distances $0, 1, \dots, d_{h,\min}$. The number of such points is given by [18]

$$V = \sum_{j=0}^{d_{h,\min}} \binom{L}{j} (2^{Mb} - 1)^j.$$

Associate one such ball with each codeword. The total number of L -tuples is $(2^{Mb})^L$. Then the maximum number of valid codewords separated by at least $d_{h,\min}$ is

$$A = \frac{(2^{Mb})^L}{V}.$$

Hence the maximum transmission rate rate is

$$\begin{aligned} R &\leq \frac{1}{L} \log_2 A \\ &= Mb - \frac{1}{L} \log_2 V. \end{aligned}$$

■

Example Assume a turbo-STC with $M = 2$, $L = 1024$, $b = 3$ (8-PSK modulation), and $d_{h,\min} = 12$. Then according to *Theorem 2*, the maximum rate is 5.84 bits/channel use, which for PSK corresponds to the same number of bits/sec/Hz. According to [1], with constant modulation, the maximum rate for a full diversity code is 3 b/s/Hz. Later in this paper, we demonstrate a rate 4 b/s/Hz code with 8PSK modulation that achieves full diversity. ■

We conclude this section on turbo-STC with TVLT with a brief discussion on implementation complexity. At the transmitter, one extra matrix multiplication is needed every symbol interval. The added complexity is not significant, especially when the number of transmit antennas is two. Also at the receiver, an extra matrix multiplication is required every symbol interval compared to codes without TVLT.

IV. DESIGN AND CONVERGENCE ANALYSIS

Conventional space time codes with full diversity and high coding gain have to be either hand designed [1] or found by exhaustive search [19]. This approach is very inefficient for turbo-STC, since two optimal component encoders do not necessarily guarantee a good overall turbo code, and the search also has to include the random interleaver. In the previous section, we have shown that turbo-STC with TVLT has full diversity and also a high probability of coding gain. This serves as a good starting point in the design of turbo-STC, but we also need an analysis tool for these codes to facilitate "tweaking" the design. Such a tool exists for binary [11], non-binary [12], and coded modulation [13], in the form of the EXIT chart. Here, we extend the EXIT chart method to the turbo space-time code with TVLT and use it to predict the convergence behavior of the turbo decoder. This extension enables to greatly expedite the code design process.

We will show that with TVLT, the code performance depends on the M singular values of the channel matrix \mathbf{H} rather than the channel matrix itself. In the rest of the paper, we exclusively consider codes for two transmit antennas. For this case, we will show that with $M = 2$, the EXIT chart can be applied to find the convergence region defined on the 2-D plane of the

channel singular values. Finally, we will show how the chart can be applied to predict the code performance.

A. Encoder

Generally speaking, space-time codes are multidimensional codes [20], [21] with the additional constraint of the rank criterion. With TVLT, full diversity is guaranteed without appealing to the rank criterion. We design the turbo-STC using accepted design practice for turbo codes. In Fig. 2, is shown the general structure of a turbo space-time code encoder with TVLT. The component encoders are two identical, recursive, systematic convolutional (RSC) encoders. The output parity bits of the two RSCs are alternately punctured. The mapping from the coded bits to symbols is the 4-D set partition proposed in [21].

Several new turbo codes are introduced in this paper for which the component encoders are listed in Table I. The notation used in the table and Fig. 2 follows [20]. The quantities m and v denote the number of input information bits and the constraint length, respectively. Each circle in Fig. 2 represents a switch. The binary variable next to each switch, $\{h_i^j, i = 1, \dots, v, j = 0, \dots, m\}$ or $\{f^k, k = 1, \dots, v\}$, indicates whether the switch is in the "on" position. For compact representation, we use the polynomial notation $H^j(D)$ to indicate the connections of the input bit j to the mod 2 adders of the convolutional code,

$$H^j(D) = \sum_{i=1}^v h_i^j D^i.$$

The feedback is represented by $H^0(D)$ and the feedforward by $F(D)$,

$$F(D) = \sum_{k=1}^v f^k(D).$$

All numbers shown in the table are octal. For example, for Code 3, the $H^3(D)$ coefficients are $(04)_8 = (001000)_2$. Then

$$H^3(D) = 1 \cdot D^3 + 0 \cdot D^2 + 0 \cdot D^1 + 0 \cdot D^0,$$

or, $h_1^3 = 0$, $h_2^3 = 0$, $h_3^3 = 0$, and $h_4^3 = 1$.

B. Decoder

The structure of the turbo space-time decoder is shown in Fig. 3. It is based on [2] and [22]. It consists of two MAP decoders, which exchange extrinsic information. The major difference with the binary turbo decoder is that the information passed from one component decoder to another may be a mixture of extrinsic and systematic information, as shown next. In binary turbo decoders, the information exchanged between component decoders contains only extrinsic information.

Let \mathbf{L}_p and \mathbf{L}_a denote respectively, the *a posteriori* and *a priori* information in log-likelihood form. Consider any one of the component decoders. Then there are two possible cases: (i) at time t the parity bits from the corresponding component encoder are transmitted; (ii) at time interval t the parity bits from the encoder are punctured. Then the *a posteriori* information can be expressed

$$\mathbf{L}_p(t) = \begin{cases} \mathbf{L}_a(t) + \mathbf{L}_{es}(t), & \text{(i)} \\ \mathbf{L}_a(t) + \mathbf{L}_e(t), & \text{(ii)} \end{cases}, \quad (8)$$

where $\mathbf{L}_e(t)$ denotes the extrinsic information and $\mathbf{L}_{es}(t)$ denote the mixture of extrinsic and systematic (intrinsic) information. The expression for the *a priori* information is:

$$\mathbf{L}_a(t) = \begin{cases} \mathbf{L}_{es}(t), & \text{(ii)} \\ \mathbf{L}_e(t), & \text{(i)} \end{cases}. \quad (9)$$

For more details, see the reference above. These differences with binary turbo decoder lead to differences in the generation of the EXIT charts for turbo-STC as discussed in Section IV-D.

C. TVLT Channel

In this subsection, we discuss the effect of TVLT on the channel model (1). The singular value decomposition (SVD) of the channel gain matrix \mathbf{H} is $\mathbf{H} = \mathbf{U}\mathbf{\Sigma}\mathbf{W}^\dagger$, where \mathbf{U} is unitary $N \times N$ and for, $M = 2$, \mathbf{W} is unitary 2×2 . The diagonal matrix $\mathbf{\Sigma}$ contains the singular values of \mathbf{H} . With that, the channel model (1) becomes

$$\mathbf{r}(t) = \mathbf{U}\mathbf{\Sigma}\mathbf{W}^\dagger\mathbf{G}(t)\mathbf{c}(t) + \mathbf{n}(t) \quad (10)$$

or

$$\mathbf{U}^\dagger\mathbf{r}(t) = \mathbf{\Sigma}\mathbf{W}^\dagger\mathbf{G}(t)\mathbf{c}(t) + \mathbf{U}^\dagger\mathbf{n}(t). \quad (11)$$

Denote $\mathbf{y}(t) = \mathbf{U}^\dagger \mathbf{r}(t)$, and $\mathbf{Q}(t) = \mathbf{W}^\dagger \mathbf{G}(t)$, with $\mathbf{Q}(t)$ a random unitary matrix that changes each time interval. Also, let the noise term $\boldsymbol{\eta}(t) = \mathbf{U}^\dagger \mathbf{n}(t)$. Then the equivalent channel model is

$$\mathbf{y}(t) = \boldsymbol{\Sigma} \mathbf{Q}(t) \mathbf{c}(t) + \boldsymbol{\eta}(t). \quad (12)$$

Since $\mathbf{Q}(t)$ changes randomly, the equivalent channel model is a "forced fast fading" channel. When the transmitted codeword is infinitely long, it will experience all possible values of $\mathbf{Q}(t)$. Consequently, the channel state is parameterized only by $\boldsymbol{\Sigma}$, or equivalently by the two singular values σ_1 and σ_2 . If the codeword is long enough, the performance of a code will be invariant over different channels, which have the same singular values.

By applying TVLT, the number of parameters needed to characterize a $2 \times N$ channel has been reduced to one or two real numbers. In contrast, without TVLT, the channel is described by $2N$ complex numbers. This reduction makes it practical to apply to EXIT chart to the analysis of turbo-STC as shown next.

D. EXIT Chart

The EXIT chart method [11], [23], [24], [12] is a convenient, low complexity tool to analyze the convergence of iterative decoding and predict performance of turbo codes. The turbo decoder consists of two component maximum *a posteriori* (MAP) decoders, which exchange extrinsic information as the iteration proceeds. Convergence to the correct decisions depends on whether each component decoder can generate better extrinsic information than it receives from the other decoder. One assumption of this method is that the codeword length or interleaver size are long enough such that the extrinsic information log-likelihood ratios (LLRs) for different time intervals are independent [11].

The EXIT chart is a graphical representation of the input/output relation of a decoder obtained through simulations. Since usually the two component decoders are the same, only one decoder needs to be tested. To perform the test, one needs to feed the component decoder with the *a priori* LLRs of the transmitted data symbols as if they are passed from the other decoder. The decoder is also fed with channel observations, generated according to a given channel model \mathcal{H} and SNR. The decoder is run and the extrinsic information is measured at its output. The EXIT chart is then the plot of the function

$$I_{out} = f(I_{in}, \mathcal{H}, \text{SNR}), \quad (13)$$

where I_{in} is the mutual information between the transmitted data symbol and the *a priori* information (in LLR form), and I_{out} is the mutual information between the transmitted data and the output extrinsic information (also in LLR form).

The most important application of EXIT chart method is to predict the performance of turbo codes. For a single antenna AWGN channel, \mathcal{H} can be described by a unit gain. Then the EXIT chart is the plot of the function

$$I_{out} = f(I_{in}, \text{SNR}). \quad (14)$$

For a given SNR, we can do tests for different I_{in} . And if we find that

$$I_{out} > I_{in}, \quad \forall I_{in} \geq 0, \quad (15)$$

then the decoder will converge to correct decisions with high probability. Otherwise, the decoding will fail. It has been shown in [11] that I_{out} is a nondecreasing function of SNR. By generating the EXIT charts (14) for increasing SNR values, a threshold SNR_{th} can be found such that if $\text{SNR} > \text{SNR}_{th}$ then (15) holds, and the decoder converges. The usefulness of SNR_{th} is to predict the beginning of the waterfall region of the bit error rate (BER).

For a single input single output (SISO), block fading channel, the channel \mathcal{H} can be modeled by a complex scalar h that follows some distribution and changes independently between codewords. The EXIT chart is the plot of the function

$$I_{out} = f(I_{in}, |h|^2 \text{SNR}).$$

The threshold SNR_{th} found from the EXIT chart in the AWGN channel can be used to predict the frame error rate (FER) over the block fading channel. The convergence region for the SISO, block fading channel is $\{h \mid |h|^2 > \text{SNR}_{th}/\text{SNR}\}$. The the frame error rate is the probability that the channel gain falls into the region where the decoding fails to converge,

$$\text{FER} = \text{Pr} \left\{ |h|^2 < \frac{\text{SNR}_{th}}{\text{SNR}} \right\}.$$

For an $M \times N$ MIMO, block fading channel, the EXIT chart is the plot of

$$I_{out} = f(I_{in}, \mathbf{H}^\dagger \mathbf{H} \text{SNR}).$$

Since \mathbf{H} is constituted from $2 \times M \times N$ real numbers, the region of convergence has to be defined in a $2 \times M \times N$ -dimensional space. Even for $M = 2$, $N = 1$, searching the threshold

boundary of a four-dimensional space is impractical. This fact impeded the development of EXIT charts for turbo space-time codes. As explained in Section , application of TVLT solves this problem. The MIMO block fading channel is transformed into an equivalent forced fast fading channel, which can be parameterized by the eigenvalues of $\mathbf{H}^\dagger \mathbf{H}$ SNR (or the singular values of the channel gain matrix \mathbf{H}). For $M = 2$, there are at most two eigenvalues. That is, the EXIT chart is the plot of the function

$$I_{out} = f(I_{in}, \lambda_1, \lambda_2),$$

where $\lambda_{1,2}$ are the eigenvalues of $\mathbf{H}^\dagger \mathbf{H}$ SNR. We will use EXIT chart to find the region of convergence on the 2-D plane defined by the two eigenvalues. The values that lie outside the region of convergence can be used to predict the FER performance of turbo space-time codes with TVLT. Examples will be shown in section IV-E.

To generate the EXIT chart for turbo-STC with TVLT, we take an approach similar to the one we proposed in [13] for SISO turbo trellis coded modulation. The input to the component MAP decoder is either generated as pure extrinsic information or synthesized as a mixture of extrinsic and systematic information (see (9)). Also, the output of the component MAP decoder is either pure extrinsic information or a mixture of extrinsic and systematic information (8). In the latter case, the systematic information needs to be removed before the extrinsic information is measured. The synthesis and separation of extrinsic information and systematic information are the major differences between EXIT charts for binary and space-time turbo codes.

Fig. 4 provides examples of EXIT charts of turbo-STC with TVLT for three fixed and different channel realizations, but with the same pair of channel eigenvalues, $\lambda_1 = 1.5$, $\lambda_2 = 1.0$. The code represented is Code 1 in Table I. In this figure, I_{in} and I_{out} , measured in bits-per-channel-use, label the abscissa and ordinate, respectively. The three curves are plots of $I_{out} = f(I_{in}, \mathbf{H})$ parameterized by the channel realizations. Since each transmission includes two information bits, I_{in} and I_{out} are in the range [0,2] bits. From the figure we can see the three EXIT curves almost overlap with each other.

E. Performance Analysis

The EXIT chart is generated for analyzing the convergence of turbo-STC with TVLT and for evaluating the FER performance over fading channel.

1) *Convergence Analysis*: For the convergence analysis we need the following theorem, based on the optimality of the constituent MAP decoders.

Theorem 3: Let I_{out} denote the mutual information between the transmitted data symbol and the output extrinsic information from a MAP decoder. Also let I_{in} denote the mutual information between the transmitted data symbol and the input *a priori* information to the MAP decoder. Consider the channel model (10) with $M = 2$ and $\Sigma = \text{diag}(\sqrt{\lambda_1}, \sqrt{\lambda_2})$, where $\text{diag}(\cdot)$ denotes a diagonal matrix. Then $I_{out} = f(I_{in}, \lambda_1, \lambda_2)$ is a nondecreasing function of both λ_1 and λ_2 .

Proof: The channel model (10) can be regarded as an equivalent AWGN channel with symbols transformed by $\mathbf{U}\Sigma\mathbf{W}^\dagger\mathbf{G}(t)$. We will first show that in this AWGN channel, the Euclidean distance between symbols is an increasing function of the eigenvalues λ_1, λ_2 . Without loss of generality, consider a scenario in which only λ_1 has been increased to $\lambda_1 + \Delta$, $\Delta > 0$. Let $d^2(\lambda_1, \lambda_2)$ and $d^2(\lambda_1 + \Delta, \lambda_2)$ denote the respective Euclidean distances. Then for $\forall \mathbf{U}, \mathbf{W}, \mathbf{c}_1(t), \mathbf{c}_2(t), \lambda_1$, and λ_2 ,

$$\begin{aligned} d^2(\lambda_1 + \Delta, \lambda_2) &\triangleq \left\| \mathbf{U} \text{diag}(\sqrt{\lambda_1 + \Delta}, \sqrt{\lambda_2}) \mathbf{W}^\dagger \mathbf{G}(t) (\mathbf{c}_1(t) - \mathbf{c}_2(t)) \right\|^2 \\ &= \left\| \mathbf{U} \text{diag}(\sqrt{\lambda_1}, \sqrt{\lambda_2}) \mathbf{W}^\dagger \mathbf{G}(t) (\mathbf{c}_1(t) - \mathbf{c}_2(t)) \right\|^2 \\ &\quad + \left\| \mathbf{U} \text{diag}(\sqrt{\Delta}, 0) \mathbf{W}^\dagger \mathbf{G}(t) (\mathbf{c}_1(t) - \mathbf{c}_2(t)) \right\|^2 \\ &> \left\| \mathbf{U} \Sigma_1 \mathbf{W}^\dagger \mathbf{G}(t) (\mathbf{c}_1(t) - \mathbf{c}_2(t)) \right\|^2 \\ &= d^2(\lambda_1, \lambda_2). \end{aligned}$$

For the scenario with the higher eigenvalue, the MAP decoder produces the lower bit error rate (BER). If a smaller eigenvalue was needed for increasing the mutual information, then the two scenarios can be made equivalent by injecting noise to match the increase in Euclidean distance. The injected noise would, however, increase the BER, which contradicts the optimality of the MAP decoder. Thus, MAP optimality dictates that the extrinsic information is non-decreasing with the eigenvalues. ■

Hence, given a λ_1 , there is a minimum λ_2 , denoted $\lambda_{2,\min}(\lambda_1)$, such that if $\lambda_2 > \lambda_{2,\min}(\lambda_1)$, then $I_{out} > I_{in}, \forall I_{in} \geq 0$, which guarantees the convergence of the decoder. The value $\lambda_{2,\min}(\lambda_1)$ is found by gradually increasing λ_2 from a small value, until I_{out} climbs above the diagonal line. An example for Code 1 is shown in Fig. 5. For $\lambda_1 = 1$, $\lambda_{2,\min}(1)$ is found to be 1.2. Repeating the process of setting the value of λ_1 and evaluating $\lambda_{2,\min}(\lambda_1)$ by the EXIT chart, results in

a λ_1 - λ_2 characteristic. In Fig. 6, there are shown λ_1 - λ_2 characteristics of the three codes listed in Table I. The codes have rates of 2, 3, and 4 b/s/Hz, respectively. For each code, the area above the curve is the region of convergence. The higher the data rate of a code, the smaller the convergence region. In addition to displaying the region of convergence, the λ_1 - λ_2 characteristic provides a convenient way to compare and select turbo space-time codes with TVLT, by quickly predicting the FER performance of the different codes, as shown in the next.

2) *FER Analysis:* The λ_1 - λ_2 characteristics described previously are applied to evaluate the FER of turbo-STC with TVLT. For a $2 \times N$ Rayleigh fading channel, where $N \geq 2$, the joint distribution of the eigenvalues of $\mathbf{H}^\dagger \mathbf{H}$ SNR is given by [25]

$$p(\lambda_1, \lambda_2) = \frac{1}{2 \text{SNR}^2} \exp\left(-\frac{1}{\text{SNR}^3} (\lambda_1 + \lambda_2) (\lambda_1 - \lambda_2)^2\right). \quad (16)$$

Denote the region of non-convergence Γ . The FER of the code at a given signal to noise ratio SNR is the integration of $p(\lambda_1, \lambda_2)$ over the region of non-convergence such as shown in Fig. 6. That is,

$$\text{FER} = \iint_{\Gamma} p(\lambda_1, \lambda_2) d\lambda_1 d\lambda_2. \quad (17)$$

The integration in (17) can be solved either analytically or numerically. The accuracy of (17) will be demonstrated in the next section by comparison to simulation results.

An analytical expression of the FER can be found for the 2×1 channel. In this case, $\mathbf{H}^\dagger \mathbf{H}$ SNR has a single eigenvalue. Let $\mathbf{H} = [h_1 \ h_2]^T$, then $\lambda = (|h_1|^2 + |h_2|^2) \text{SNR}$. For Rayleigh fading, the eigenvalue λ follows a chi-square distribution with four degrees of freedom

$$p(\lambda) = \frac{1}{2 \text{SNR}^2} \lambda e^{-\lambda/\text{SNR}}. \quad (18)$$

The minimum value of λ for which the code converges is the point $(0, \lambda_{2,\min})$ in Fig. 6. We denote this value λ_{th} , and from the figure find for Code 1, $\lambda_{th} \simeq 3.8$. Then the FER is expressed

$$\begin{aligned} \text{FER} &= \int_0^{\lambda_{th}} p(\lambda) d\lambda \\ &= \int_0^{\lambda_{th}} \frac{4}{\gamma^2} \lambda^2 \exp\left(-\frac{2\lambda^2}{\gamma}\right) d\lambda^2 \\ &= 1 - \left(\frac{\lambda_{th}}{\text{SNR}} + 1\right) \exp\left(-\frac{\lambda_{th}}{\text{SNR}}\right). \end{aligned} \quad (19)$$

V. NUMERICAL RESULTS

In this section, numerical results are provided for the new codes with TVLT (Table I) and their performance is compared to the case when no TVLT is applied. The interleaver length is 1024 for all codes. The performance is measured in terms of the FER versus SNR.

The performance of Code 1, with a spectral efficiency of 2 b/s/Hz with 4-PSK modulation, is shown in Fig. 7, for both 2×1 and 2×2 configurations and with and without TVLT. It is observed that with TVLT, the code has a steeper FER decline with SNR, and both 2×1 and 2×2 configurations are visually parallel with the outage probability curves. Without TVLT, the codes display a marked loss in diversity gain in both configurations. The turbo-STC with TVLT gap to outage capacity is about 1-1.5 dB. This example demonstrates the ability of TVLT to provide full diversity gains (as discussed in earlier sections). The FER curve obtained from the λ_1 - λ_2 characteristic is also shown in the figure. It is observed that the prediction fits the simulation sufficiently well that can be used as a quick tool to evaluate performance of turbo space-time codes.

The performance of the 8-PSK, 3 b/s/Hz Code 2 is presented in Fig. 8 for cases similar to those in Fig. 7. It is observed that Code 2 loses diversity gain in the absence of TVLT, but the loss is not as marked as in the case of Code 1. A possible explanation is that the 8-PSK modulation here achieves a better rank of the error matrix than the 4-PSK modulation of Code 1. In fact, this observation reinforces the argument for TVLT, which achieves the rank criterion through diversity of the modulation. Back to the figure, it is observed that the FER predicted from the EXIT chart is reasonably close match with the simulation results.

The last in the series of performance analyses is that of Code 3 displayed in Fig. 9. This code achieves a spectral efficiency of 4 b/s/Hz with 8-PSK modulation. As such, it breaks the rate limit of full diversity space-time codes with fixed modulation, which for 8-PSK is 3 b/s/Hz [1]. Conversely, codes at 4 b/s/Hz cannot achieve full diversity. This is clearly evident for the code without TVLT in Fig. 9. With TVLT however, full diversity is regained, as predicted by *Theorem 2* in Section III. The EXIT chart predicts performance within 1 dB of the simulation results.

VI. CONCLUSIONS

In this paper time-varying linear transformations are applied to designing full diversity turbo space-time codes. TVLT consists of a unitary transformation which converts conventional constant modulations into time varying modulations practically guaranteeing that the rank criterion is satisfied. Unlike space-time and turbo space-time codes with fixed modulations, which may require exhaustive rank tests to ensure full diversity, the rank property is inherent to TVLT codes. Rules for the construction of TVLT have been provided for an arbitrary number of antennas. The relation between full diversity and coding gain has been formalized in a theorem, which states that the coding gain fails to meet a prescribed level with a probability that decreases exponentially with the minimum Hamming distance of the code. This theorem has two implications: (1) that the design criterion for turbo-STC is the Hamming distance between codewords (before TVLT application); (2) since turbo-STC typically display high minimum Hamming distance, they are good candidates for use with TVLT. A second theorem established the rates achievable with TVLT space-time codes. It was shown that limits due to constant modulation over the codeword duration do not apply to TVLT and that significantly higher rates are achievable.

Beyond affording full diversity gain, robust coding gains, and high rates, TVLT also enables convenient performance analysis and design using the EXIT method. Previously, the EXIT chart analysis has been developed for binary and multilevel codes, but it has been limited to SISO systems. We have shown that in generating the EXIT chart for turbo codes with TVLT, the performance is averaged over channels specified by the singular values of the channel matrix. In particular, for the two transmit antenna case, two eigenvalues are sufficient to specify the region of convergence of the turbo code. It was also shown that this region of convergence can be applied to evaluate the FER of the code. Subsequently, the EXIT chart was found to be a useful tool in the design and analysis of turbo space-time codes.

Finally, various turbo space-time codes with TVLT were demonstrated by simulations and by FER analysis. These codes display all the properties of full diversity, gain and rate as predicted by the analysis and overall have excellent performance with a gap of only 2 dB to outage capacity.

APPENDIX

A. *Proof of* $\det(\overline{\mathbf{E}}\overline{\mathbf{E}}^\dagger) \geq \det(\mathbf{V}(\mu)\mathbf{V}(\mu)^\dagger)$

Proof: Matrix $\overline{\mathbf{E}}$ incorporates $\mathbf{V}(\mu)$ and has additional columns, denoted $\phi_1, \dots, \phi_{d_h-\mu}$. According to the rank-one updates theorem for determinants [26], $\det(\mathbf{A} + \mathbf{c}\mathbf{d}^\dagger) = \det(\mathbf{A})(1 + \mathbf{d}^\dagger\mathbf{A}^{-1}\mathbf{c})$ where \mathbf{A} is $n \times n$ and \mathbf{c}, \mathbf{d} are $n \times 1$. Then,

$$\det(\overline{\mathbf{E}}\overline{\mathbf{E}}^\dagger) = \det\left(\mathbf{V}(\mu)\mathbf{V}(\mu)^\dagger + \sum_{i=1}^{d_h-\mu} \phi_i\phi_i^\dagger\right).$$

We apply the theorem one column at a time

$$\begin{aligned} \det(\overline{\mathbf{E}}\overline{\mathbf{E}}^\dagger) &= \det\left(\left(\mathbf{V}(\mu)\mathbf{V}(\mu)^\dagger + \sum_{i=2}^{d_h-\mu} \phi_i\phi_i^\dagger\right) + \phi_1\phi_1^\dagger\right) \\ &= \det\left(\mathbf{V}(\mu)\mathbf{V}(\mu)^\dagger + \sum_{i=2}^{d_h-\mu} \phi_i\phi_i^\dagger\right) (1 + \phi_1^\dagger\phi_1). \end{aligned}$$

Repeating this step for all columns, we prove the claim

$$\begin{aligned} \det(\overline{\mathbf{E}}\overline{\mathbf{E}}^\dagger) &= \det(\mathbf{V}(\mu)\mathbf{V}(\mu)^\dagger) \prod_{i=1}^{d_h-\mu} (1 + \phi_i\phi_i^\dagger) \\ &\geq \det(\mathbf{V}(\mu)\mathbf{V}(\mu)^\dagger). \end{aligned}$$

■

REFERENCES

- [1] V. Tarokh and A. Seshadri, "Space-time codes for high data rate wireless communication: Performance criterion and code construction," *IEEE Trans. on Info. Theory*, vol. 44, pp. 744–765, March 1998.
- [2] D. Cui and A. Haimovich, "Performance of parallel concatenated space-time codes," *IEEE Commun. Lett.*, vol. 5, pp. 236–238, June 2001.
- [3] Y. Liu, P. Fitz, and O. Y. Takeshita, "Full rate space-time turbo codes," *IEEE J. on Select. Areas Commun.*, vol. 19, pp. 969–980, May 2001.
- [4] A. Stefanov and T. M. Duman, "Turbo-coded modulation for systems with transmit and receive antenna diversity over block fading channels: system model, decoding approaches, and practical considerations," *IEEE J. on Select. Areas Commun.*, vol. 19, pp. 958–968, May 2001.
- [5] A. Wittneben, "A new bandwidth efficient transmit antenna modulation diversity scheme for linear digital modulation," *In Proc. IEEE ICC' 93*, vol. 3, pp. 1630 – 1634, May 1993.
- [6] A. Hiroike, F. Adachi, and N. Nakajima, "Combined effects of phase sweeping transmitter diversity and channel coding," *IEEE Trans. on Veh. Technol.*, vol. 41, pp. 170 –176, May 1992.
- [7] B. Su and S. Wilson, "Phase sweeping transmitter diversity in mobile communications," *in Proc. IEEE VTC '96*, vol. 1, pp. 131 – 135, April 1996.

- [8] X. Ma and G. B. Giannakis, "Space-time-multipath coding using digital phase sweeping," in *Proc. IEEE GLOBECOM '02*, vol. 1, pp. 384 – 388, Nov 2002.
- [9] A. Gutierrez, J. Li, S. Baines, and D. Bevan, "An introduction to PSTD for is-95 and cdma2000," *IEEE Wireless Communications and Networking Conference, WCNC'99*, vol. 3, pp. 1358 – 1362, Sept. 1999.
- [10] W. Shi, C. Kominakis, R. Wesel, and B. Daneshrad, "Robustness of space-time turbo codes," in *Proc. IEEE ICC '01*, vol. 6, pp. 11–14, June 2001.
- [11] S. T. Brink, "Convergence behavior of iterative decoded parallel concatenated codes," *IEEE Trans. on Commun.*, vol. 49, pp. 1727–1737, Oct. 2001.
- [12] A. Grant, "Convergence of non-binary iterative decoding," in *Proc. IEEE GLOBECOM '01*, vol. 2, pp. 1058–1062, Nov 2001.
- [13] H. Chen and A. Haimovich, "Exit charts for turbo trellis coded modulation," *To appear on the November issue of IEEE Commun. Lett.*, 2004.
- [14] L. C. Perez, J. Seghers, and D. J. Costello Jr., "A distance spectrum interpretation of turbo codes," *IEEE Trans. on Info. Theory*, vol. 42, pp. 1698–1709, Nov. 1996.
- [15] A. Perotti and S. Benedetto, "A new upper bound on the minimum distance of turbo codes," *to appear in IEEE Trans. on Info. Theory*. available at <http://commgroup.polito.it/personal/perotti/pubs/dmin-ub.pdf>.
- [16] M. Breiling, S. Peeters, and J. Huber, "Interleaver design using backtracking and spreading methods," in *Proc. Int. Symp. Info. Theory*, p. 451, June 2000.
- [17] O. Y. Takeshita and D. J. Costello Jr., "New deterministic interleaver designs for turbo codes," *IEEE Trans. on Info. Theory*, vol. 51, pp. 1988 – 2006, Sept. 2000.
- [18] S. G. Wilson, "Digital modulation and coding," *Prentice-Hall, Inc.*, pp. 432–434, 1996.
- [19] S. Baro and G. Bauch, "Improved codes for space-time trellis-coded modulation," *IEEE Commun. Lett.*, vol. 4, pp. 20–22, Jan. 2000.
- [20] G. Ungerboeck, "Channel coding with multilevel/phase signals," *IEEE Trans. on Info. Theory*, vol. 28, pp. 55–67, Jan. 1982.
- [21] S. S. Pietrobon, R. H. Deng, A. Lafanechere, G. Ungerboeck, and D. J. Costello Jr., "Trellis-coded multidimensional phase modulation," *IEEE Trans. on Info. Theory*, vol. 36, pp. 63–89, Jan 1990.
- [22] P. Robertson and T. Worz, "Bandwidth-efficient turbo trellis-coded modulation using punctured component codes," *IEEE J. on Select. Areas Commun.*, vol. 16, pp. 206–218, Feb 1998.
- [23] H. E. Gamal and A. R. Hammons, "Analyzing the turbo decoder using Gaussian approximation," *IEEE Trans. on Info. Theory*, vol. 47, pp. 671–686, Feb. 2001.
- [24] D. Divsalar, S. Donlinar, and F. Pollara, "Iterative turbo decoder analysis based on density evolution," *IEEE J. on Select. Areas Commun.*, vol. 19, pp. 891–907, May 2001.
- [25] M. Chiani, M. Z. Win, A. Zanella, and J. H. Winters, "Exact symbol error probability for optimum combining in the presence of multiple cochannel interferers and thermal noise," in *Proc. IEEE GLOBECOM '01*, vol. 2, pp. 1182–1186, 2001.
- [26] C. D. Meyer and C. Meyer, "Matrix analysis and applied linear algebra," *Society for Industrial and Applied Math*, p. 457, February 2001.

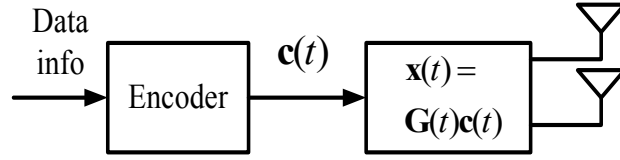


Fig. 1. Turbo space-time encoder with TVLT.

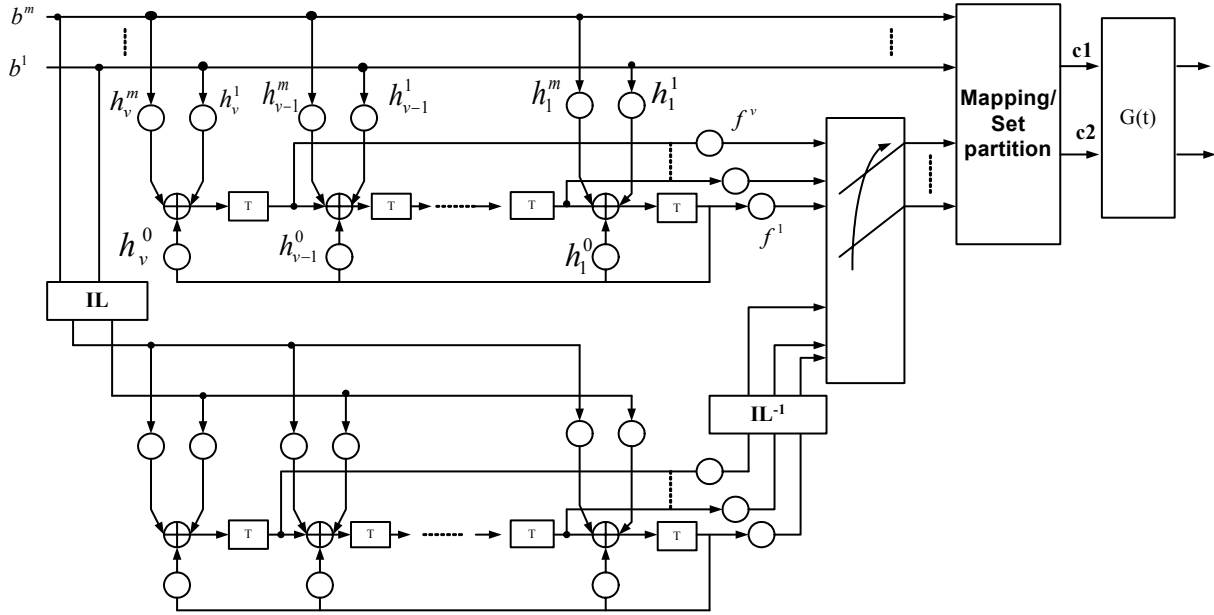


Fig. 2. The turbo-STC encoder. The encoder is constituted from recursive, systematic, convolutional component codes.

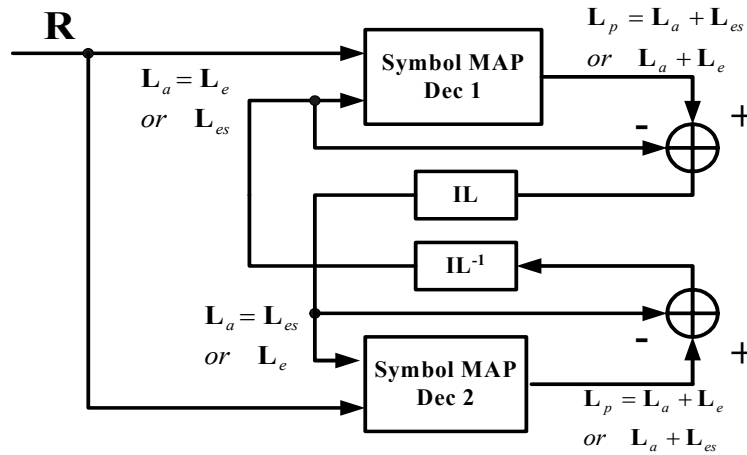


Fig. 3. Structure of the turbo space-time decoder.

TABLE I

LIST OF ENCODERS USED IN THIS PAPER.

Code	m	v	$H^0(D)$	$H^1(D)$	$H^2(D)$	$H^3(D)$	$H^4(D)$	$F(D)$
(1)	2	3	6_8	1_8	2_8			3_8
(2)	3	4	10_8	01_8	02_8	04_8		07_8
(3)	4	4	10_8	01_8	02_8	04_8	10_8	03_8

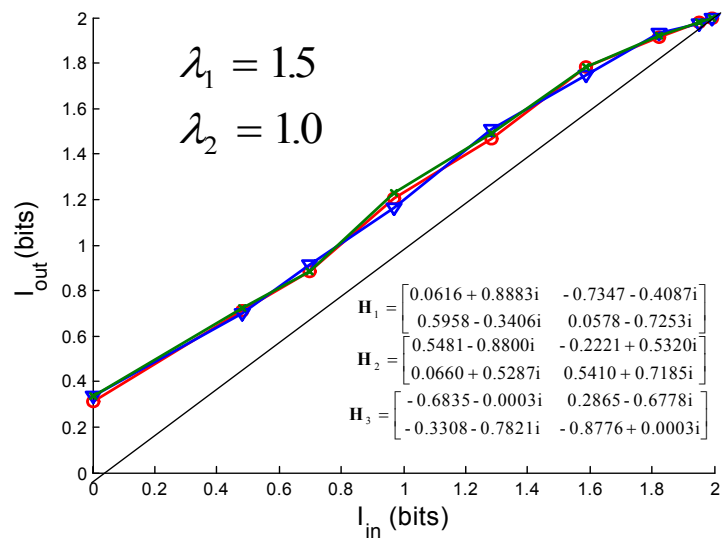


Fig. 4. Demonstration: turbo-STC has same convergence behavior over different channels with the same eigenvalues.

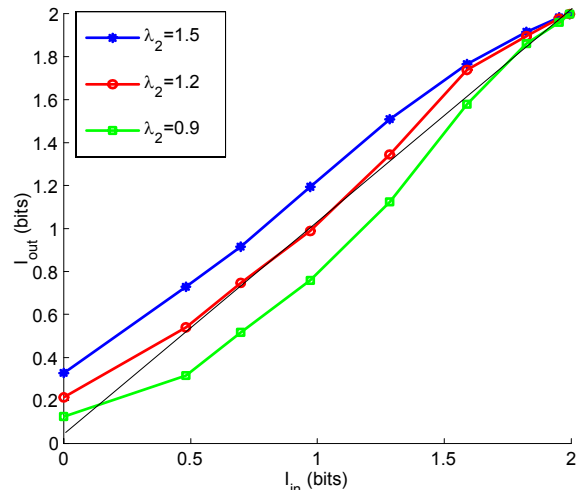


Fig. 5. The EXIT charts of the 2 b/s/Hz code for eigenvalues $\lambda_1 = 1$ and $\lambda_2 = 0.9, 1.2, 1.5$.

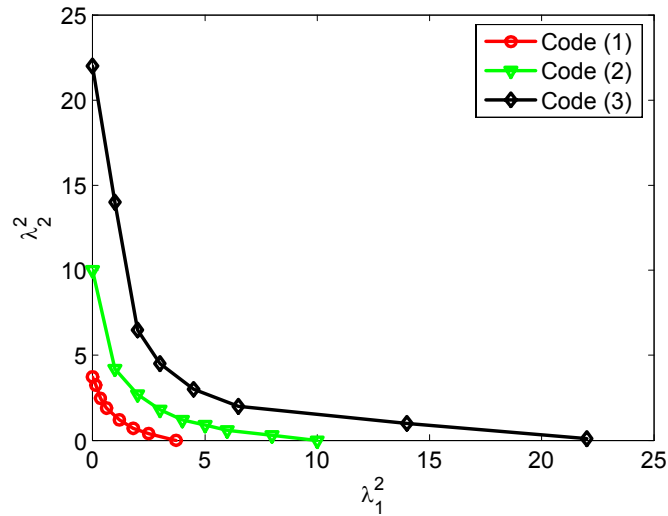


Fig. 6. Convergence regions of the proposed codes.

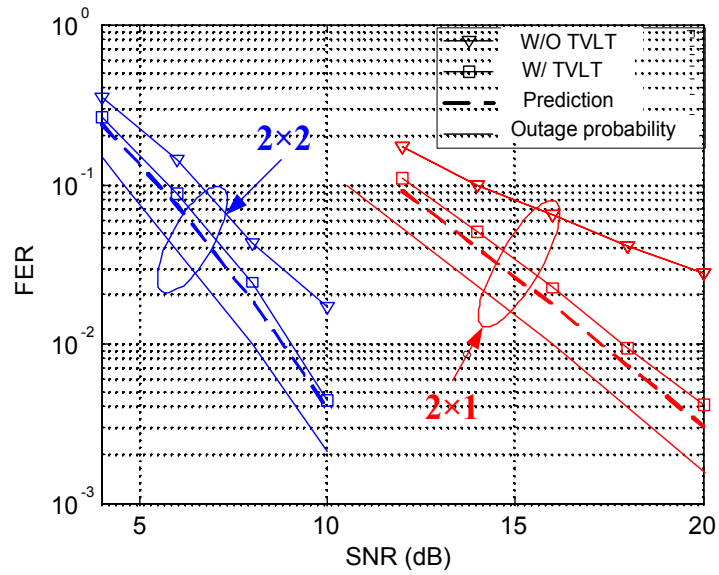


Fig. 7. FER performance comparison of the 2 b/s/Hz Code 1 with and without TVLT.

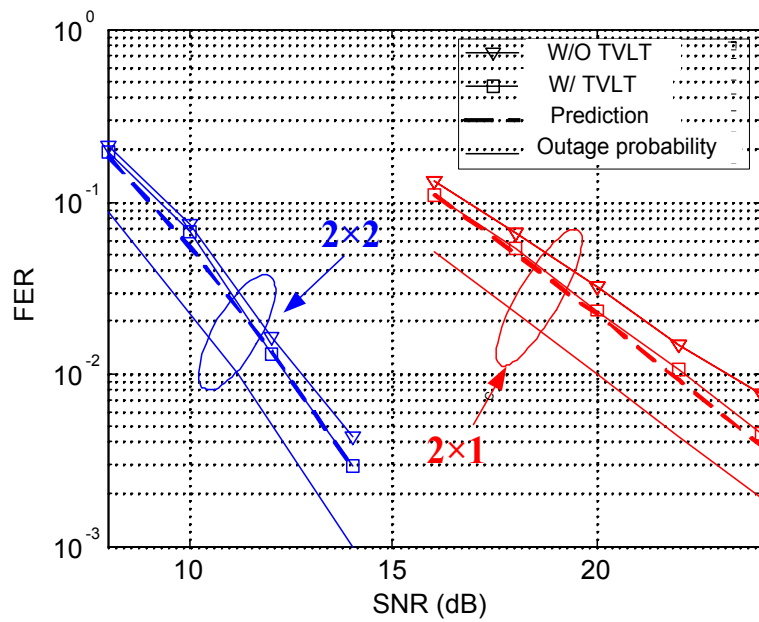


Fig. 8. FER performance comparison of the 3 b/s/Hz Code 2 with and without TVLT.

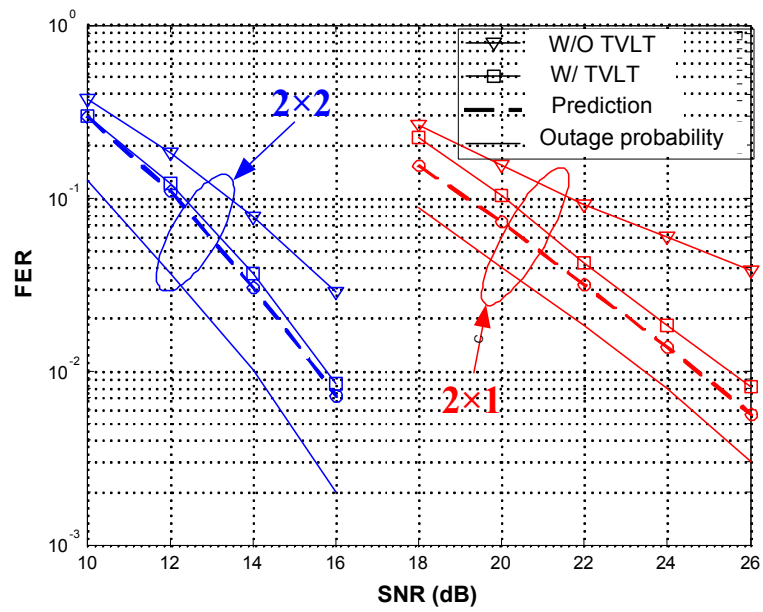


Fig. 9. FER performance comparison of the 4 b/s/Hz Code 3 with and without TVLT.