Performance Analysis and Constituent Code Design for Space-Time Turbo Coded Modulation over Fading Channels

Djordje Tujkovic, Ian Oppermann
Centre for Wireless Communications (CWC), University of Oulu
P. O. Box 4500, FIN-90014 Oulu, Finland
email:djordje@ee.oulu.fi, ian@ee.oulu.fi

(*Invited presentation*)

Abstract

Recent advances in space-time coding and array processing have lead to substantial improvements in broadband wireless access systems. A design method for constituent recursive space-time trellis codes (Rec-STTrCs) and parallel concatenated space-time turbo coded modulation (STTuCM) was introduced recently as a new framework for building low complexity, large equivalent constraint-length STTrCs. In this paper, STTuCM is analyzed in terms of the two-dimensional distance spectrum (DS), the truncated union bound (UB) and the iterative decoding convergence. The derived bounds appeared to be tight over fast fading channels but rather optimistic over quasi-static fading channels. Nevertheless, the rank deficient part of the DS has been well enumerated, which turned out to be crucial for the accurate design of the full transmit diversity STTuCM in the second half of the paper.

Modified design criteria for space-time (ST) codes over quasi-static and fast fading channels were further introduced to capture the joint effects of error exponents and multiplicities in the code's DS. A recursive systematic form with a primitive equivalent feed-backward polynomial was assumed for constituent codes (CCs) to assure good convergence of iterative decoding. The DS optimization produced new sets of CCs for STTuCM resulting in a robust ST coding scheme with improved performance over both quasi-static and fast fading channels. Full transmit diversity was achieved with no constrains on the implemented pseudo-random bit-wise odd-even information interleaving.

Index Terms

Space-time codes, turbo coded modulation, union bound, convergence analysis, code design.

I. Introduction

RECENT results in information theory [1], [2] have lead to a new way of dealing with the inherent problems of communication for broadband wireless access. New field measurement campaigns [3], [4] have demonstrated the potentially large capacity increase for multiple-input multiple-output (MIMO) wireless channels [5], [6]. Building on space-time (ST) processing techniques [7], ST coding [8], [9] approaches the MIMO channel capacity by simultaneously encorrelating signals across spatial and temporal domains.

An integrated trellis coding and modulation design with a redundancy expansion in both the signal and antenna space, space-time trellis codes (STTrCs) [10], [11] can be characterized as a generalization of classical trellis coded modulation [12], [13] to multi-antenna systems. The high computation complexity resulting from heuristic techniques for reducing the cardinality of the class of codes has constrained STTrC optimization to low constraint-lengths, typically up to 64-state trellis codes. Typically designed to achieve the maximum transmit diversity gain, such STTrCs commonly settled with rather modest coding gains. Another disadvantage of limited freedom, optimization efforts have almost exclusively been for either slow or fast fading channels. In practical applications characterized by extremely variable Doppler frequencies, existing STTrCs fail to demonstrate much needed robustness. A recent ad-hoc attempt to build a large constraint-length, 256-state STTrC in [14] resulted in a highly non-optimized brute force solution as shown in [15].

In an analogy to single antenna turbo coded modulation from [16], a design method for constituent recursive STTrCs (Rec-STTrCs) and parallel concatenated space-time turbo coded modulation (STTuCM) was recently introduced in [17]. The new ST coding framework, applicable to any STTrC from the literature, integrated code concatenation into a random-like ST coding approach. Similar to single antenna turbo codes [18], the large equivalent constraint-lengths and randomness were jointly provided by an information interleaver, a building block that due to iterative decoding does not add considerably to decoding complexity. Paper [17] considered a punctured recursive constituent code (CC) design based on 4-, 8- and 16-state *Tarokh et al.* STTrCs (Tarokh-STTrCs) from [10], handcrafted for a spectral efficiency of 2 bps/Hz with two transmit antennas and QPSK modulation. The approach was further extended in [19], [15] to 3 and 4 bps/Hz employing 8PSK and 16QAM modulation,

This work was supported by Nokia, Elektrobit, Finnish Air Force, the National Technology Agency of Finland (Tekes) and Graduate School of Electronics, Telecommunications and Automation (GETA).

respectively. Similar schemes were studied in parallel in [20], [21], [22] and their accompanied papers [23], [24], [25], with distinct CCs and information interleaving structural differences that will become apparent after examining the system model in the next section. The transfer function based performance analysis of single antenna turbo coded modulation from [16] was recently studied over AWGN channels in [26]. Resembling the performance analysis proposed in [27] and [28] for single-antenna turbo coded modulation from [29], authors in [21] and [23] added an additional channel interleaver between encoding and modulation to enable the tractable transfer function based DS enumeration of the STTuCM over AWGN and quasi-static fading channels, respectively. However, such a radical discrepancy towards the parallel concatenated STTrC model utilized in simulations is somewhat difficult to justify. Other ST turbo code proposals including [30], [31], [32] can be distinguished by the single antenna turbo coded modulations they were inspired by [9].

The outstanding performance of the STTuCM seen through simulations in narrowband [15], [17], and broadband OFDM [33], WCDMA [34], [35] and MC-CDMA [36] based wireless systems motivated the theoretical studies conducted in this paper. The paper is organized as follows. After a short review of the system model in Section II, the main notations and definitions are given in Subsection III-A. The two configurations suggested in [17], namely the fully recursive STTuCM with both CCs being Rec-STTrCs and the non-fully recursive STTuCM where one of the CCs was replaced with an equivalent feed-forward STTrC, are comparatively studied in Section III. The performance analysis includes the unified two-dimensional distance spectrum (DS) interpretation in Subsection III-B, the truncated UB over quasi-static and fast fading channels in Subsection III-C and the iterative decoding convergence analysis in Subsection III-D.

The design of new CCs for STTuCM is addressed in Section IV. The modified design criteria for ST codes over quasi-static and fast fading channels, which capture the joint effects of error exponents and multiplicities in the code's DS, are introduced in Subsection IV-A. A recursive systematic form with a primitive equivalent feed-backward polynomial is adopted in Subsection IV-B to define a subset of candidate CCs for the computer search. The DS optimization in Subsection IV-C produced new sets of CCs for STTuCM resulting in a robust ST coding scheme with improved performance over both quasi-static and fast fading channels as shown in Subsection IV-D. The main results and conclusions driven throughout the paper are summarized in Section V.

II. System Model

A. Encoder

Consider a system with N transmit and M receive antennas depicted in Fig. 1. As in [15], [17], the transmitter employs a parallel concatenation of two CCs followed by puncturing, multiplexing and channel interleaving. The input information frame $\boldsymbol{x} = [\boldsymbol{x}(1), \ \boldsymbol{x}(2), \ldots, \boldsymbol{x}(L)]$ consisting at each time instant $l, l = 1, 2, \ldots, L$, of a sequence of K bits $\boldsymbol{x}(l) = [x_1(l), x_2(l), \ldots, x_K(l)]$ is first encoded by the first CC to produce a coded frame $\boldsymbol{c}^1 = \left[\boldsymbol{c}^1(1), \ \boldsymbol{c}^1(2), \ldots, \boldsymbol{c}^1(L)\right]$. At each time instant l, a coded frame consists of N complex symbols $\boldsymbol{c}^1(l) = \left[c_1^1(l), c_2^1(l), \ldots, c_N^1(l)\right]$ each belonging to a 2^Z complex constellation with unit average energy. The interleaved version of the input information sequence is then encoded by the second CC to produce a coded frame $\boldsymbol{c}^2 = \left[\boldsymbol{c}^2(1), \boldsymbol{c}^2(2), \ldots, \boldsymbol{c}^2(L)\right]$, $\boldsymbol{c}^2(l) = \left[c_1^2(l), c_2^2(l), \ldots, c_N^2(l)\right]$. To preserve the overall rate, the two coded frames are punctured and multiplexed so that at one time instant, only one of the component encoders has access to N transmit antennas. Unlike [15], [17], only the trellis of the first CC was terminated at the end of the transmission frame while the trellis of the second CC was left open. The resulting STTuCM coded frame is denoted by $\boldsymbol{c} = [\boldsymbol{c}(1), \boldsymbol{c}(2), \ldots, \boldsymbol{c}(L)]$, $\boldsymbol{c}(l) = [c_1(l), c_2(l), \ldots, c_N(l)]$ with $\boldsymbol{c}(l) = \boldsymbol{c}^1(l)$ for l odd and $\boldsymbol{c}(l) = \boldsymbol{c}^2(l)$ for l even. Symbol $c_n(l)$, $n = 1, 2, \ldots, N$ is assigned to transmit antenna n during time instant l. Prior to transmission, symbol streams on transmit antennas are interleaved with the same symbol-wise channel interleaver π_c.

B. Information Interleaver

Apart from distinct CCs, the information interleaver structure is the main feature that distinguishes different STTuCMs proposed in parallel in [17], [20], [21], [22]. In [20] a single pseudo-random bit-wise interleaver of length LK was applied and no puncturing was assumed resulting in STTuCM with a reduced rate as compared to its constituent STTrCs. In [21] a single symbol-wise pseudo-random interleaver of length L was utilized. Due to additional symbol de-interleaving employed after the second CC, the input information bits were equally protected. However with no constrains imposed on the information interlever, the second CC was non-uniformly punctured. In [22] the same symbol-wise odd–even information interleaver from [16] was adopted. Like the one implemented in [21], such an interleaver assures that the positions of the input bits in the input symbol $\boldsymbol{x}(l) = [x_1(l), x_2(l), \ldots, x_K(l)]$ remain unchanged. The interleaver utilized in [15], [17] and in this paper consists of two pseudo-random bit-wise interleavers, π_{O} for odd l and π_{E} for even l input positions, i.e., it enables the mapping of any input bit from a given even (odd) input symbol to any bit position of any even (odd) symbol.

C. Decoder

The block diagram of the STTuCM decoder is depicted in Figure 1(b). Prior to iterative decoding, channel de-interleaving and received signal de-multiplexing is undertaken. The component codes are decoded with symbol-by-symbol MAP decoders. Due to

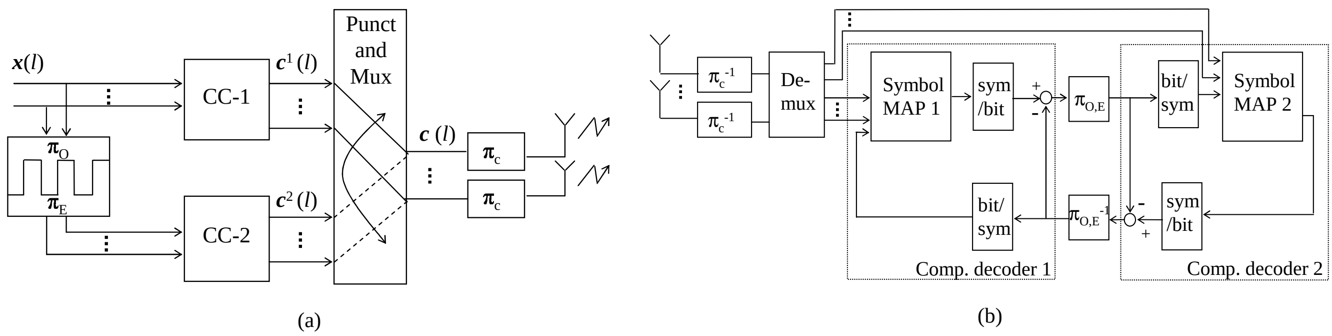


Fig. 1. System model. (a) Parallel concatenated encoder. (b) Iterative decoder.

the assumed bit-wise information interleaving, component decoders include additional symbol-to-bit and bit-to-symbol reliability transforms. Hence, the exchange of extrinsic information between the two component decoders during iterative decoding is done on the bit level. When symbol-wise information interleaving is employed, the exchange of extrinsic information is done directly on the symbol level, i.e., the schemes proposed in [21], [22] operate on the symbol level. The superior performance of the bit-level STTuCM from [17] over the symbol-level STTuCM from [21] was demonstrated in [15]. For more details on STTuCM decoder design issues see [15], [17].

III. Performance Analysis

A. Upper Bounds over Fading Channels

Assuming maximum likelihood (ML) decoding with a perfect knowledge of channel state information (CSI) at the receiver, the UB on the frame error rate (FER) for a given ST code over fading channels with AWGN is given as

$$P_{\mathrm{FER}} = \sum_{\mathcal{D}} H\left(\Delta_{\mathrm{H}}, \Delta_{\mathrm{P}}\right) P_{\mathrm{PEP}}\left(\Delta_{\mathrm{H}}, \Delta_{\mathrm{P}}\right) \tag{1}$$

where $\Delta_{\mathrm{H}} = \Delta_{\mathrm{H}}\left(\boldsymbol{c}, \boldsymbol{e}\right)$ and $\Delta_{\mathrm{P}} = \Delta_{\mathrm{P}}\left(\boldsymbol{c}, \boldsymbol{e}\right)$ denote the effective Hamming distance (EHD) and the effective product distance (EPD) between the transmitted sequence $\boldsymbol{c}$ and its erroneously decoded counterpart $\boldsymbol{e}$. The set $\mathcal{D}$ includes all distinct $\langle \Delta_{\mathrm{H}}, \Delta_{\mathrm{P}} \rangle$ pairs in a two-dimensional code's DS. The EHD and EPD are commonly referred to as the error coefficients or error exponents. $H\left(\Delta_{\mathrm{H}}, \Delta_{\mathrm{P}}\right)$s in (1) are multiplicities in a two dimensional code's DS, i.e., the average number of error events $\{\boldsymbol{c} \rightarrow \boldsymbol{e}\}$ with a given EHD and EPD. The second term in the product from (1) is the pair-wise error probability (PEP) as a function of error exponents. Based on the exact polynomial PEP function, a new simplified PEP bound that is tighter than the standard Chernoff bound in [10] and asymptotically tight with high signal-to-noise ratios (SNRs) has been derived in [37] as

$$P_{\mathrm{PEP}}\left(\Delta_{\mathrm{H}}, \Delta_{\mathrm{P}}\right) \leq \binom{2\Delta_{\mathrm{H}} M - 1}{\Delta_{\mathrm{H}} M - 1} \Delta_{\mathrm{P}}^{-\Delta_{\mathrm{H}} M} \left(\frac{\mathrm{SNR}}{N}\right)^{-\Delta_{\mathrm{H}} M}. \tag{2}$$

In [15], a block fading channel model was adopted with the number of independent fading realizations per transmission frame B taking values from $B \in \{1, 2, 4, L\}$. For $B = 1$, the channel is denoted as quasi-static or slow, while the $B = L$ case, referred to as fast fading, represents the artificial case of an ideally channel interleaved system. The block fading model is in general suitable for fading channels in which a certain block of adjacent transmitted symbols are affected by highly correlated fading path gains. The length of the block may be considered as a first approximation of the channel's coherence time for single-carrier systems or the channel's coherence bandwidth in multi-carrier systems. On quasi-static fading channels, $\Delta_{\mathrm{P}}\left(\boldsymbol{c}, \boldsymbol{e}\right)$ is defined as the geometric mean of non-zero eigenvalues of the $N \times N$ epoch-by-epoch calculable squared code-word difference matrix (SCDM)

$$\boldsymbol{A}\left(\boldsymbol{c}, \boldsymbol{e}\right) = \sum_{l=1}^{L} \left[\Delta_1^{ec}\left(l\right), \ldots, \Delta_N^{ec}\left(l\right)\right]^{\mathrm{H}} \left[\Delta_1^{ec}\left(l\right), \ldots, \Delta_N^{ec}\left(l\right)\right] \tag{3}$$

with $\Delta_n^{ec}\left(l\right) = e_n\left(l\right) - c_n\left(l\right)$. $\Delta_{\mathrm{H}}\left(\boldsymbol{c}, \boldsymbol{e}\right)$ is given as a rank of $\boldsymbol{A}\left(\boldsymbol{c}, \boldsymbol{e}\right)$. Therefore on quasi-static fading channels, the EHD is bounded by $1 \leq \Delta_{\mathrm{H}} \leq N$. Notice that when $\Delta_{\mathrm{H}}\left(\boldsymbol{c}, \boldsymbol{e}\right) = N$, the EPD is equal to a determinant of the SCDM, i.e., $\Delta_{\mathrm{P}}\left(\boldsymbol{c}, \boldsymbol{e}\right) = \det\left(\boldsymbol{A}\left(\boldsymbol{c}, \boldsymbol{e}\right)\right)$. Likewise, on fast fading channels $\Delta_{\mathrm{P}}\left(\boldsymbol{c}, \boldsymbol{e}\right)$ is given as the geometric mean of non-zero epoch-by-epoch calculable terms

$$\lambda_l\left(\boldsymbol{c}, \boldsymbol{e}\right) = \left[\Delta_1^{ec}\left(l\right), \ldots, \Delta_N^{ec}\left(l\right)\right] \left[\Delta_1^{ec}\left(l\right), \ldots, \Delta_N^{ec}\left(l\right)\right]^{\mathrm{H}} \tag{4}$$

and $\Delta_{\mathrm{H}}\left(\boldsymbol{c}, \boldsymbol{e}\right)$ denotes the cardinality of a set of time instances $\{l\}$ in which terms in (4) are non-zero. On AWGN channels, the EPD is proportional to the squared Euclidean distance between $\boldsymbol{c}$ and $\boldsymbol{e}$, i.e., $\Delta_{\mathrm{P}}\left(\boldsymbol{c}, \boldsymbol{e}\right) = M \mathrm{trace}\left(\boldsymbol{A}\left(\boldsymbol{c}, \boldsymbol{e}\right)\right)$. The factor M

can be seen as a gain from repetition coding. The EHD is reduced to $1/M$ accordingly to accommodate the lack of diversity over AWGN channels to the unified PEP representation in (2). Alternatively, the EPD over AWGN channels can be directly calculated based on

$$\Delta_{\mathrm{P}}\left(\boldsymbol{c},\boldsymbol{e}\right)=M\sum_{l=1}^{L}\sum_{n=1}^{N}\left|\Delta_{n}^{ec}\left(l\right)\right|^{2} \tag{5}$$

which is more favorable as it consumes less memory during an epoch-by-epoch DS enumeration.

B. Distance Spectrum Interpretation

The truncated UB analysis in Subsection III-C as well as the DS optimization in Subsection IV-C rely on the two-dimensional DS interpretation of STTuCM wrt the EHD and EPD evaluated in this subsection. Without loss of generality, consider the two single pair-wise error events $\{\boldsymbol{c}^1\rightarrow\boldsymbol{e}^1\}$ and $\{\boldsymbol{c}^2\rightarrow\boldsymbol{e}^2\}$ in the first and second CC, respectively as depicted in Fig. 2. Due to its additive property, the SCDM from (3) can be decomposed [9], [38]

$$\boldsymbol{A}\left(\boldsymbol{c},\boldsymbol{e}\right)=\boldsymbol{A}\left(\boldsymbol{c}^1,\boldsymbol{e}^1\right)+\boldsymbol{A}\left(\boldsymbol{c}^2,\boldsymbol{e}^2\right) \tag{6}$$

with separated terms contributed from error events in the two CCs. On fast fading channels, as in [9], [38], the EHD and EPD can be directly decomposed following from (4)

$$\Delta_{\mathrm{P}}\left(\boldsymbol{c},\boldsymbol{e}\right)=\left(\Delta_{\mathrm{P}}\left(\boldsymbol{c}^1,\boldsymbol{e}^1\right)\right)^{\frac{\Delta_{\mathrm{H}}\left(\boldsymbol{c}^1,\boldsymbol{e}^1\right)}{\Delta_{\mathrm{H}}\left(\boldsymbol{c},\boldsymbol{e}\right)}}\left(\Delta_{\mathrm{P}}\left(\boldsymbol{c}^2,\boldsymbol{e}^2\right)\right)^{\frac{\Delta_{\mathrm{H}}\left(\boldsymbol{c}^2,\boldsymbol{e}^2\right)}{\Delta_{\mathrm{H}}\left(\boldsymbol{c},\boldsymbol{e}\right)}} \tag{7}$$

$$\Delta_{\mathrm{H}}\left(\boldsymbol{c},\boldsymbol{e}\right)=\Delta_{\mathrm{H}}\left(\boldsymbol{c}^1,\boldsymbol{e}^1\right)+\Delta_{\mathrm{H}}\left(\boldsymbol{c}^2,\boldsymbol{e}^2\right). \tag{8}$$

On AWGN channels, a spectral interpretation reduces to a one-dimensional DS wrt the EPD. From (5) it directly follows [38]

$$\Delta_{\mathrm{P}}\left(\boldsymbol{c},\boldsymbol{e}\right)=\Delta_{\mathrm{P}}\left(\boldsymbol{c}^1,\boldsymbol{e}^1\right)+\Delta_{\mathrm{P}}\left(\boldsymbol{c}^2,\boldsymbol{e}^2\right). \tag{9}$$

For an error event $\{\boldsymbol{c}\rightarrow\boldsymbol{e}\}$, let $\mathcal{M}=\mathcal{M}\left(\boldsymbol{c},\boldsymbol{e}\right)$ denote the subset of relevant PEP measures for a given type of channel. That is for fast fading channels, $\mathcal{M}$ includes the EPD and EHD while for quasi-static fading and AWGN channels, $\mathcal{M}$ is equal to the SCDF and EPD, respectively. Also, let the binary operator "$\uplus$" denote the subset of appropriate operations from (6) – (9) for a given type of channel and related performance measure. That is, for fast fading channels "$\uplus$" processes the EPDs and the EHDs of its operands according to (7) and (8), respectively. Likewise, on quasi-static fading and AWGN channels, the corresponding measures are summed according to (6) and (9), respectively. Let $\mathcal{M}^0$ denote the identity element from $\{\mathcal{M}\}$, i.e., $\mathcal{M}=\mathcal{M}\uplus\mathcal{M}^0$.

Let $\boldsymbol{x}$ and $\hat{\boldsymbol{x}}$ denote the correctly and erroneously decoded information sequences related to $\boldsymbol{c}^1$ and $\boldsymbol{e}^1$, respectively. In the example from Fig. 2, $\boldsymbol{x}=[01\,11\,00\,10\,11\,01]$ and $\hat{\boldsymbol{x}}=[01\,10\,10\,11\,11\,01]$. In addition to set $\mathcal{M}$, each error event in the given CC will be identified Hamming distance vector $\boldsymbol{h}=\boldsymbol{h}\left(\boldsymbol{c}^1,\boldsymbol{e}^1\right)$ between $\boldsymbol{x}$ and $\hat{\boldsymbol{x}}$. The two CCs share the same information error event $\{\boldsymbol{x}\rightarrow\hat{\boldsymbol{x}}\}$. That is, correctly and erroneously decoded information sequences for $\{\boldsymbol{c}^2\rightarrow\boldsymbol{e}^2\}$ are interleaved versions of $\boldsymbol{x}$ and $\hat{\boldsymbol{x}}$, respectively. Therefore, $\boldsymbol{h}\left(\boldsymbol{c}^1,\boldsymbol{e}^1\right)=\boldsymbol{h}\left(\boldsymbol{c}^2,\boldsymbol{e}^2\right)$ holds. The four components of $\boldsymbol{h}=[h_{\mathrm{O}1},h_{\mathrm{O}2},h_{\mathrm{E}1},h_{\mathrm{E}2}]$ enumerate the number of bit positions of type $\{x_k\left(l\right),\hat{x}_k\left(l\right)\}=\{0,1\}$ and type $\{x_k\left(l\right),\hat{x}_k\left(l\right)\}=\{1,0\}$ in the two information interleavers. For rigorous definition of $\boldsymbol{h}$ see [38]. For the example pair-wise error sequence from Fig. 2, $\boldsymbol{h}=[1\,0\,1\,1]$. Naturally, the sum over the components of $\boldsymbol{h}$, denoted hereafter by h, is equal to the Hamming distance between $\boldsymbol{x}$ and $\hat{\boldsymbol{x}}$, i,e., $h=3$.

The two single error events from Fig. 2 diverge from an arbitrary and possibly different trellis state and time instant l, have in general the different lengths ψ and reemerge to an arbitrary and possibly different state. A single error event, which is punctured for the first time immediately after diverging, will be denoted as odd-punctured. Similarly, a single error event, which is punctured for the first time one trellis step after diverging, will be denoted as even-punctured. In the example from Fig. 2, $\{\boldsymbol{c}^1\rightarrow\boldsymbol{e}^1\}$ is the length $\psi=3$ odd-punctured error event while $\{\boldsymbol{c}^2\rightarrow\boldsymbol{e}^2\}$ is the length $\psi=2$ even-punctured error event.

Let $H_{1,\psi}^{j}\left(\boldsymbol{h},\mathcal{M}^j\right)$ denote the average number of single error events of length ψ in the CC-j, $j=1,2$, with the given measures $\boldsymbol{h}$ and $\mathcal{M}^j$. Multiplicities $H_{1,\psi}^{j}\left(\boldsymbol{h},\mathcal{M}^j\right)$ include both the odd- and even-punctured single error events that diverge from an arbitrary state at the time instant $l=1$. To enumerate the set $H_{1,\psi}^{j}\left(\boldsymbol{h},\mathcal{M}^j\right)$ in each of the punctured CCs, the General algorithm (GA) presented in [13] for the general class of GNU codes has been modified as in the [38] to include the enumeration of both the Hamming distance vector $\boldsymbol{h}$ and the set of performance measures $\mathcal{M}^j$. Since some of the best STTrCs are geometrically non-uniform (GNU) wrt the EPD [11], the application of the standard transfer function bound to STTrCs, as proposed in [21], [23], [39] is not easily justifiable. The truncated UB for the GNU stand-alone STTrCs over AWGN and fading channels was evaluated in parallel in [40] applying the GA from [13] and in [41] through the simplified code transfer function from [42].

Unlike the stand-alone STTrCs where compounded error events can be expurgated from the UB in (2) [43], the concatenation of two or more single error events in each CC has to be accounted for when evaluating the DS of any turbo-like code [44].

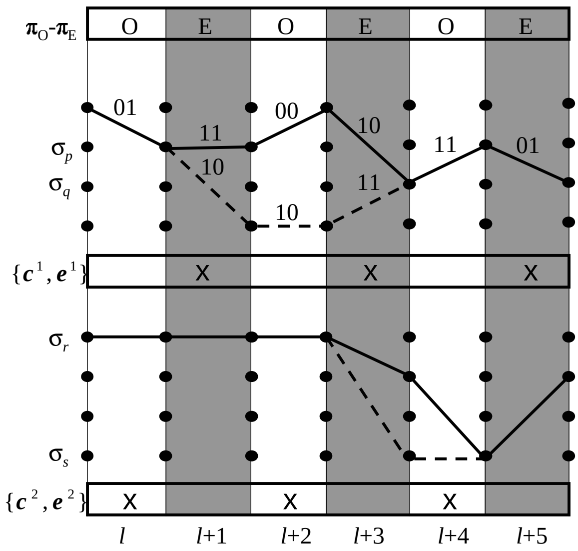


Fig. 2. Pair-wise error events in the two constituent codes and their relation to odd and even bit-wise interleavers and puncturing patterns. Branch labels denote the input information bits for $K = 2$. Symbol X means that the output signal vector is punctured.

The average number of compounded error events with the given $\boldsymbol{h}$ and $\mathcal{M}^j$ that start at an arbitrary time instant $l \geq 1$ and consist of $\Theta \geq 1$ concatenated single error events of total length $\Psi \leq L$ is determined as

$$H^j_{\Theta,\Psi}\left(\boldsymbol{h},\mathcal{M}^j\right) = \Gamma\left(\Theta,\Psi\right)\prod_{\theta=1}^{\Theta} H^j_{1,\psi_\theta}\left(\boldsymbol{h}_\theta,\mathcal{M}^j_\theta\right) \tag{10}$$

with $\boldsymbol{h} = \boldsymbol{h}_1 + \ldots + \boldsymbol{h}_\theta + \ldots + \boldsymbol{h}_\Theta$, $\Psi = \psi_1 + \ldots + \psi_\theta + \ldots + \psi_\Theta$, $\mathcal{M}^j = \mathcal{M}^j_1 \uplus \ldots \uplus \mathcal{M}^j_\theta \uplus \ldots \uplus \mathcal{M}^j_\Theta$, where $\boldsymbol{h}_\theta$, ψ_θ and $\mathcal{M}^j_\theta$ correspond to the θth error event in the concatenation. Similar to [44], the function $\Gamma\left(\Theta,\Psi\right)$ is defined as

$$\Gamma\left(\Theta,\Psi\right) = \left(\begin{array}{c} \frac{L}{2} - \lceil\frac{\Psi}{2}\rceil + \Theta \\ \Theta \end{array}\right). \tag{11}$$

The contribution of single error events that diverge at time instant $l > 1$ is also encountered with (11). The multiplicities in the two-dimensional DS of the CC-j are calculated as

$$H^j\left(\boldsymbol{h},\mathcal{M}^j\right) = \sum_{\Theta}\sum_{\Psi} H^j_{\Theta,\Psi}\left(\boldsymbol{h},\mathcal{M}^j\right). \tag{12}$$

To enable the analysis tractable, both the odd and even bit-wise information interleavers were assumed to be uniform [44], i.e., all pseudo-random permutations are equally probable. The resultant DS of STTuCM wrt $\mathcal{M}$ as averaged over all possible information interleavers is given by [38]

$$H\left(\mathcal{M} = \mathcal{M}^1 \uplus \mathcal{M}^2\right) = \sum_{\boldsymbol{h}} \frac{H^1\left(\boldsymbol{h},\mathcal{M}^1\right)H^2\left(\boldsymbol{h},\mathcal{M}^2\right)}{\left(\begin{array}{c} LK/2 \\ h_{\mathrm{O1}}, h_{\mathrm{O2}} \end{array}\right)\left(\begin{array}{c} LK/2 \\ h_{\mathrm{E1}}, h_{\mathrm{E2}} \end{array}\right)}. \tag{13}$$

For fast fading and AWGN channels, the DS wrt $\mathcal{M}$ can be directly plugged into (1). Otherwise, for the quasi-static fading case where $\mathcal{M} = \boldsymbol{A}\left(\boldsymbol{c},\boldsymbol{e}\right)$, the rank and the geometric mean of the non-zero eigenvalues have to be calculated for each SCDM in (13) to arrive at the two-dimensional DS wrt the EHD and EPD.

C. Truncated Union Bound

Let us truncate the UB in (1) by replacing the set $\mathcal{D}$ with its subset $\mathcal{P}$, i.e., *significant* part of the code's DS. The choice of parameters that uniquely determine the significant part of the DS is a complex multi-dimensional problem. The main bottleneck of the introduced DS interpretation is not its computational complexity but the memory requirements [38].

Given the memory constrains of the GA, for the quasi-static fading case all punctured single error events with a Hamming distance up to $h_{\max} = 6$, a sequence length up to $\psi_{\max} = 6$ and an EPD up to $\Delta_{\mathrm{P}\max} = 16$ were first enumerated in each of the CCs. Compounded error events consisting of up to $\Theta_{\max} = 3$ concatenated single error events were then formed. On fast fading channels, the significant part of the DS was identified with $h_{\max} = 6$, $\psi_{\max} = 6$, $\Delta_{\mathrm{P}\max} = 48$ and $\Theta_{\max} = 3$. Despite the large increase in the number of DS entries, a further increase in any of the above parameters for 8-state CC from [17] over fading channels has been empirically found to insignificantly contribute to the sum in (1).

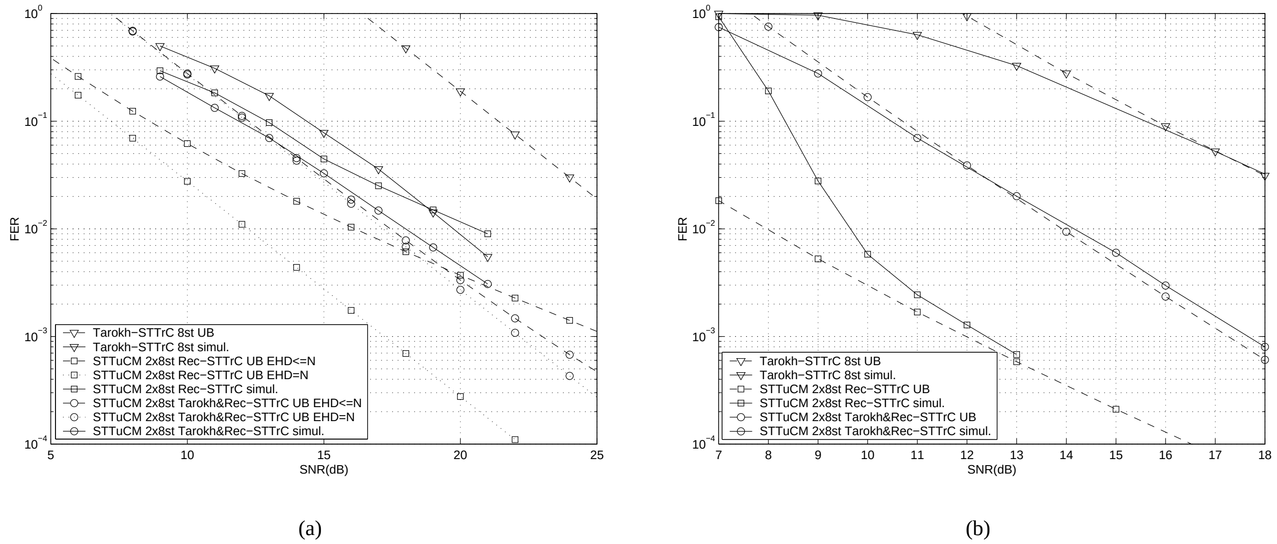


Fig. 3. The simulation and the union bound performance of the fully recursive and the non-fully recursive STTuCM, 8-state CCs, 10 decoding iterations, $LK = 260$ bits, $N = 2$, $M = 1$. (a) Quasi-static fading channel. (b) Fast fading channel.

Fig. 3 depicts the FER analytical and simulation results of the STTuCM over quasi-static and fast fading channels. The two configurations proposed in [15], [17] and applied throughout papers [33], [34], [35] were comparatively studied. The first one denoted “STTuCM 2x8st Rec-STTrC” was built from two 8-state Rec-STTrCs from [17]. In the second configuration denoted “STTuCM 2x8st Tarokh&Rec-STTrC”, one of the CCs was replaced with a non-recursive Tarokh-STTrC [10]. Such a configuration was shown in [17] to lead to better performance over quasi-static fading channels. Notice that both configurations employ the non-systematic CCs. Both STTuCMs achieve a spectral efficiency of 2 bps/Hz applying QPSK modulation ($Z = K = 2$) over $N = 2$ transmit and $M = 1$ receive antennas. The length of the information frame was set to $LK = 260$ bits and 10 decoding iterations were assumed in simulations.

With no additional bounding of the conditional PEP as in [45], [40], upper bounds of the STTuCM over quasi-static fading channels in Fig. 3(a) can be characterized as rather optimistic. Nevertheless, the rank deficient error events seem to have been well enumerated, which appeared to be crucial for the accurate design of the full rank STTuCM in Section IV. The STTuCM exhibited the effect of spectral thinning [46], which resulted in a DS with very few codewords with low EHDs and EPDs, and a majority of codewords with average-to-high EHDs and EPDs, similar to the DS expected from random-like ST codes. Evidently, the determinant criterion [10], commonly applied for design of the stand-alone STTrCs may not necessarily be adequate for CC optimization in the STTuCM framework.

From Fig. 3(a), it is apparent that without particular CC and/or information interleaving optimization, the punctured parallel concatenation in general does not preserve the full rank of its CCs. That is, the two-dimensional DS of STTuCM over quasi-static fading channels has the non-zero entries $H(\Delta_{\mathrm{H}}, \Delta_{\mathrm{P}})$ for $\Delta_{\mathrm{H}} < N$. For the fully recursive configuration, a quite large “code energy” concentrated in the rank deficient part of the DS severely deteriorated the performance. On the other hand, the non-fully recursive configuration experienced a negligible performance deviation away from the second order diversity slope in the range of low-to-medium SNRs typically exhibited in practice.

On fast fading channels, UB of STTuCM in Fig. 3(b) appeared to be tight. The similar tightness was observed for constituent non-punctured stand-alone STTrCs. The configuration with both Rec-STTrCs exhibited the stronger water-falling effect, i.e., the higher slope in the transition region towards the “error-floor”. The error-floor here refers to the asymptotical high SNR behavior with the slope of the FER curve characterized by the diversity gain. The non-fully recursive STTuCM experienced no performance degradation in the low SNR region, i.e., the water-fall and error-floor regions were indistinguishable. The reasons for such interesting behavior will become apparent after examining the iterative decoding convergence properties in the next subsection.

D. Iterative Decoding Convergence

Similar to [47], [48], an i.i.d. symmetric Gaussian approximation of extrinsic information was assumed in this paper to trace the convergence of iterative decoding. As in [38], each component decoder from Fig. 1(b) was modelled as a non-linear system with a certain transfer function. The SNR at the input of the jth component decoder, denoted hereafter by $\mathrm{SNR}^{\mathrm{i}}_{\mathrm{d}j}$, was varied and the corresponding output SNR, denoted hereafter by $\mathrm{SNR}^{\mathrm{o}}_{\mathrm{d}j}$, measured to empirically determine the non-linear transfer function G_j. Functions $\mathrm{SNR}^{\mathrm{o}}_{\mathrm{d}1} = G_1\left(\mathrm{SNR}^{\mathrm{i}}_{\mathrm{d}1}\right)$ and $\mathrm{SNR}^{\mathrm{i}}_{\mathrm{d}2} = G_2^{-1}\left(\mathrm{SNR}^{\mathrm{o}}_{\mathrm{d}2}\right)$ were plotted on the same graph. This

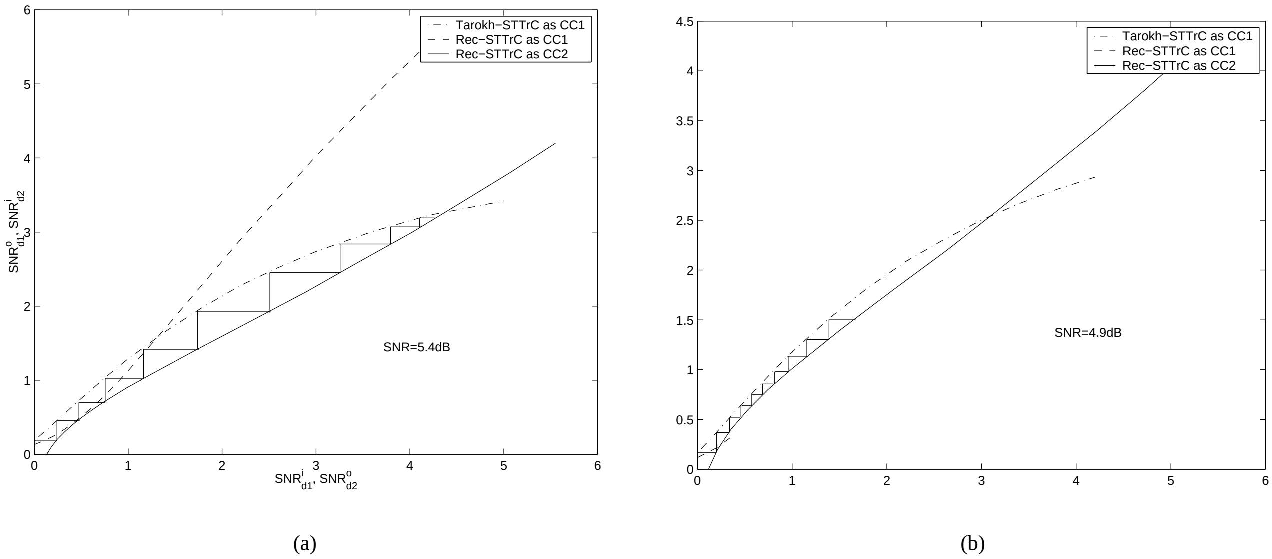


Fig. 4. Iterative decoding convergence for fully recursive and non-fully recursive STTuCM on an AWGN channel. (a) SNR=5.4dB. (b) SNR=4.9dB.

was done comparatively in Fig. 4 for two STTuCMs from Subsection III-C assuming an AWGN channel with SNRs of 5.4 dB and 4.9 dB. The convergence threshold was defined in [47] as the channel SNR value at which two transfer functions just touch. Above this SNR value, the "tunnel" of iterative decoding opens [47] and the $\mathrm{SNR}^{\mathrm{o}}_{\mathrm{d}j}$ increases over iterations without bound. In practise, this means that the probability of error, inversely proportional to such $\mathrm{SNR}^{\mathrm{o}}_{\mathrm{d}j}$ [48], converges towards the error-floor even with finite, large enough value of $\mathrm{SNR}^{\mathrm{o}}_{\mathrm{d}j}$.

From Fig. 4(a), the convergence threshold of the fully-recursive STTuCM is 5.4 dB. On the other hand, the tunnel of iterative decoding for the non-fully recursive STTuCM is never open. The first CC being non-recursive has a zero asymptotic slope of its extrinsic information transfer function, which conflicts with the asymptotic slope one of the recursive second CC. However, due to the almost complementary outlook of its transfer functions at low $\mathrm{SNR}^{\mathrm{i}}_{\mathrm{d}i}$, the non-fully recursive STTuCM enables reasonably good evolution of its extrinsic information within 10 decoding iterations. At the channel SNR=5.4 dB, the non-fully recursive STTuCM has a remarkably larger value of $\mathrm{SNR}^{\mathrm{o}}_{\mathrm{d}j}$ than the recursive one whose tunnel is pinched at a very low value of $\mathrm{SNR}^{\mathrm{o}}_{\mathrm{d}j}$. Fig. 4(b) shows that even lower values of channel SNR enable a reasonable evolution of $\mathrm{SNR}^{\mathrm{o}}_{\mathrm{d}j}$ for a non-fully recursive configuration.

The good convergence properties are essential for the satisfactory performance of turbo-like codes over quasi-static fading channels [48]. Adding to it a less deteriorating effect of the rank deficient part of the DS, it is easy to understand the superior performance of the non-fully recursive STTuCM over the fully recursive one experienced through simulations over quasi-static fading channels. On fast fading channels where good DS characteristics play a more important role, the fully-recursive STTuCM outperformed the non-fully recursive one in spite of its inferior convergence properties.

IV. Constituent Code Design

A. Design Criteria

The maximization of the minimum EHD, coined the *rank* criterion on slow fading channels and the *distance* criterion on fast fading channels was identified as the primary importance ST coding design criterion in [10]. The maximization of the minimum EPD, denoted the *determinant* criterion on slow fading channels and *product* criterion on fast fading channels was established as the secondary importance ST coding design criterion therein. In [49] it was further argued that for a moderately large diversity order, i.e., $\Delta_{\mathrm{H}} M \geq 4$, a MIMO fading channel converges reasonably close to an AWGN channel. In such a case the PEP is inversely proportional to the code squared Euclidean distance in (5). The maximization of it, denoted the *trace* criterion, was introduced in [49] as a secondary importance design criterion. The rank and the distance criteria were sustained therein as the primary ones to meet the condition $\Delta_{\mathrm{H}} M \geq 4$ with fewer receive antennas. Similar conclusions on the importance of the code squared Euclidean distance, though as a primary design criterion over quasi-static fading channels for systems with a large number of receive antennas but also those operating in the low SNR region were driven in [50], [51]. In the moderate SNR region, the PEP was further determined in [50] to be inversely proportional to $\det(\boldsymbol{I}+\boldsymbol{A})$ where $\boldsymbol{I}$ denotes the identity matrix. The equal eigenvalue criterion was introduced independently in [52], [53] in a context of the determinant criterion and optimal coding gain, arguing that the maximization of the minimum code squared Euclidean distance renders eigenvalues of the SM to be equal, which maximizes the minimum EPD.

All of the above design criteria concentrated on error exponents. As a result, different codes with the same minimum EHD and EPD sometimes exhibited different performance in simulations. To capture the joint effects of pair-wise error events, with error exponents both equal to and greater than the minimum, but also the contribution of multiplicities to the union bound in (1), the following design criteria for ST codes over quasi-static fading channels were introduced in [9], [54]:

- *Significant full rank*. To maximize the diversity gain over quasi-static fading channels, the STC must satisfy the full rank criterion in its significant part of the DS, i.e., $H\left(\Delta_{\mathrm{H}},\Delta_{\mathrm{P}}\right)=0$ for $\forall\langle\Delta_{\mathrm{H}},\Delta_{\mathrm{P}}\rangle\in\mathcal{P}|_{\Delta_{\mathrm{H}}<N}$.
- *Minimum spectral product*. To maximize the coding gain, among all ST codes that satisfy the significant full rank criterion choose the one that minimizes

$$\Delta_{\mathrm{DS}}=\sum_{\langle\Delta_{\mathrm{H}},\Delta_{\mathrm{P}}\rangle\in\mathcal{P}|_{\Delta_{\mathrm{H}}=N}} H\left(\Delta_{\mathrm{H}},\Delta_{\mathrm{P}}\right)\Delta_{\mathrm{P}}{}^{-\Delta_{\mathrm{H}}M}. \tag{14}$$

The significant full rank criterion guaranties the full spatial diversity given the large enough cardinality of set $\mathcal{P}$. Notice that (14) follows directly from (1) and (2), cancelling the terms independent of the EPD. Once the ST code's DS wrt $\mathcal{M}=\boldsymbol{A}$ in (13) is enumerated, the minimum spectral product optimization can be performed based on either of three EPD definitions, i.e., assuming $\Delta_{\mathrm{P}}=\det\left(\boldsymbol{A}\right)$, $\Delta_{\mathrm{P}}=\det\left(\boldsymbol{I}+\boldsymbol{A}\right)$ or $\Delta_{\mathrm{P}}=\mathrm{trace}\left(\boldsymbol{A}\right)$.

The importance of a code's DS, though in the context of stand-alone STTrCs and methods for its optimization over quasi-static fading channels were in parallel also studied in [55]. A recent attempt to apply the similar criteria from (14) to fast fading channels was made in [56]. However, on fast fading channels, it is not possible to accurately remove the SNR term from the DS optimization. The pair-wise error events with the same EPD may in general have different lengths and therefore experience different levels of diversity [9], [57]. Hence, given the SNR point, the DS optimization over fast fading channels in Subsection IV-C will assume the minimization of the truncated UB in (1).

B. Subset of Candidate Constituent Codes

In contrast to [15], [17] where CCs were defined by trellis diagrams, here we adopted a transform domain model with binary generator polynomials [9], [54]. To each of K parallel input bits from $\boldsymbol{x}\left(l\right)=\left[x_1\left(l\right),x_2\left(l\right),\ldots,x_K\left(l\right)\right]$, a binary recursive feedback shift register with memory order ν_k, $k=1,\ldots,K$, was associated. The sum $\nu=\nu_1+\nu_2+\ldots+\nu_K$ denoted the total memory order of the CC and each ν_k satisfied $\lfloor\nu/K\rfloor\leq\nu_k\leq\lceil\nu/K\rceil$. Each $c_n\left(l\right)$ was assumed to be the result of mapping Z coded bits $\boldsymbol{y}_n\left(l\right)=\left[y_n^1\left(l\right),y_n^2\left(l\right),\ldots,y_n^Z\left(l\right)\right]$ to a given 2^Z level amplitude and/or phase modulation. Fig. 5 depicts the kth shift register and its relation with the uth register, $u=1,\ldots,K$. Introducing an indeterminate D to denote a delay operator, the input/output relations can be represented using a binary polynomial arithmetic relation

$$y_n^z\left(D\right)=\sum_{k=1}^{K}\left(f_{nzk}^x x_k\left(D\right)+f_{nzk}\left(D\right)a_k\left(D\right)\right) \tag{15}$$

with $z=1,\ldots,Z$ and $a_k\left(D\right)$ denoting a signal entering the kth shift register

$$a_k\left(D\right)=\sum_{k=1}^{K}\left(b_{ku}^x x_u\left(D\right)+b_{ku}\left(D\right)Da_u\left(D\right)\right). \tag{16}$$

f_{nzk}^x and b_{ku}^x are binary constants, $f_{nzk}\left(D\right)$ is a degree ν_k binary feed-forward polynomial from shift register k to the output bit z at the transmit antenna n, $b_{ku}\left(D\right)$ is a degree ν_u-1 binary feed-backward polynomial from shift register u to the input of the shift register k. Assuming that CCs in parallel concatenation are the same, the optimization over the code subset defined with (15), (16) requires a tedious search over $2^{NZ(\nu+2K)+K(\nu+K)}$ different codes. In the following we restrict to the special case of $N=K=Z=2$, $\nu=3$ ($\nu_1=1,\nu_2=2$) and the natural QPSK mapping, i.e., $c_n\left(l\right)=\sqrt{E_s}\exp\{\sqrt{-1}\left(y_n^1\left(l\right)+2y_n^2\left(l\right)\right)\pi/2\}$. Therefore, the redundancy expansion is assumed only in the antenna space. The above special case covers a quite large set of 2^{38} 8-state CCs with a spectral efficiency of 2 bps/Hz. Therefore we further constrain the set of candidate codes to:

- Subset of systematic CCs since they were shown in [47] for classical single antenna turbo codes to result in better convergence of iterative decoding. Therefore the following parameters were chosen: $f_{112}^x=f_{121}^x=1$, $f_{112}\left(D\right)=f_{121}\left(D\right)=0$, $f_{122}^x=f_{111}^x=0$, $f_{122}\left(D\right)=f_{111}\left(D\right)=0$, $f_{212}^x=f_{222}^x=f_{211}^x=f_{221}^x=0$.
- Fully connected trellises without parallel transitions, i.e., $b_{11}^x=b_{22}^x=1$ and $b_{12}^x=b_{21}^x=0$ respectively.

Solving (15) and (16) with the above assumptions, the input-output relations in the CC can be further simplified as

$$y_1^z\left(D\right)=x_{(3-z)}\left(D\right) \tag{17}$$

$$y_2^z\left(D\right)=\sum_{k=1}^{2}\frac{f_{zk}^e\left(D\right)}{b^e\left(D\right)}x_k\left(D\right) \tag{18}$$

with $z=1,2$ and equivalent feed-forward and feed-backward polynomials given as

$$f_{zk}^e\left(D\right)=f_{2zk}\left(D\right)\left(1+Db_{(3-k)(3-k)}\left(D\right)\right)+Df_{2z(3-k)}\left(D\right)b_{(3-k)k}\left(D\right) \tag{19}$$

$$b^e\left(D\right)=1+\left(b_{11}\left(D\right)+b_{22}\left(D\right)\right)D+\left(b_{11}\left(D\right)b_{22}\left(D\right)+b_{12}\left(D\right)b_{21}\left(D\right)\right)D^2. \tag{20}$$

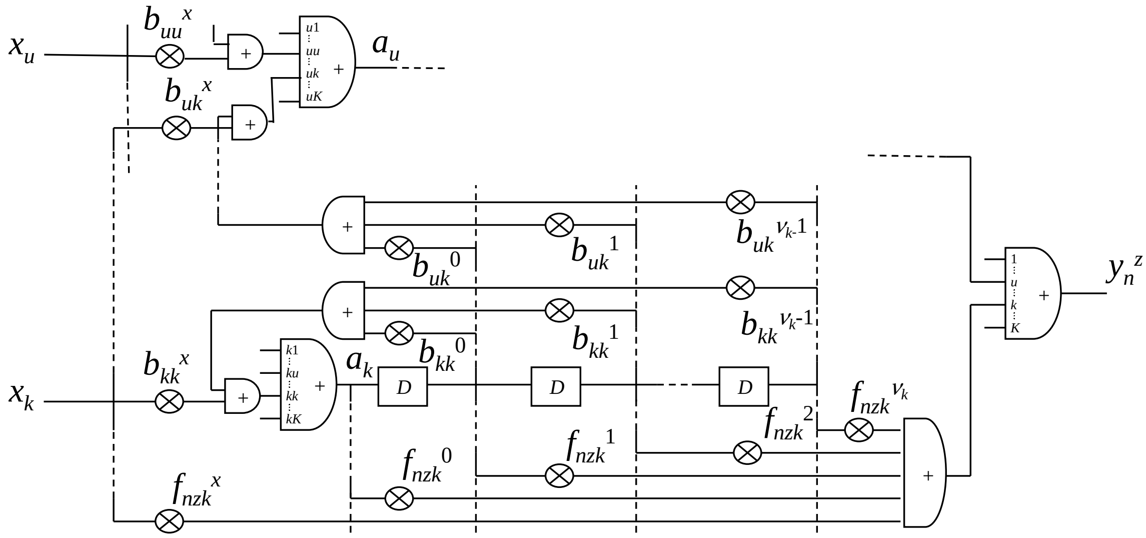


Fig. 5. The kth shift register in the constituent code.

Notice that a degree ν binary polynomial $b^e(D)$ is a function of feed-backward polynomials only.

For each k, let the polynomial $h_k(D) = x_k(D) + \hat{x}_k(D)$ denote a sequence of one step Hamming distances between $x_k(D)$ and $\hat{x}_k(D)$. Due to recursive feedback, the summed coefficients of $h_k(D)$ over all k has the minimum value of 2, which occurs in one of the four distinct cases: 1) $h_1(D) = 1 + D^{L_1}$, $h_2(D) = 0$; 2) $h_1(D) = 0$, $h_2(D) = 1 + D^{L_2}$; 3) $h_1(D) = 1$, $h_2(D) = D^{L_3}$; 4) $h_1(D) = D^{L_4}$, $h_2(D) = 1$. Following the intuition from [46], the maximization of $\min(L_1, L_2, L_3, L_4)$ is expected to result in code-word gaining EHD and EPD. Choosing the degree ν primitive polynomial for $b^e(D)$ maximizes the minimum L_1 and L_2 to $L_1 = L_2 = 2^\nu - 1 = 7$. A primitive $b^e(D)$ is also expected to result in better convergence of iterative decoding [47].

Let us define a matrix of polynomial coefficients $\boldsymbol{B}(D) = [b_{11}(D);\ b_{12}(D); b_{22}(D); b_{21}(D)]$, with $b_{ku}(D) = b_{ku}^0 + b_{ku}^1 D + \ldots + b_{ku}^{\nu_u - 1} D^{\nu_u - 1}$, $k, u = 1, 2$. Letting $b^e(D) = 1 + D + D^3$ results in four solutions for the feed-backward polynomials, i.e., $\boldsymbol{B}_1 = [10; 10; 01; 10]$, $\boldsymbol{B}_2 = [10; 01; 00;\ 10]$, $\boldsymbol{B}_3 = [00; 11; 11; 10]$ and $\boldsymbol{B}_4 = [00; 01;\ 10; 10]$. $\boldsymbol{B}_1$ and $\boldsymbol{B}_2$ have $\min(L_3, L_4) = 2$ while $\boldsymbol{B}_3$ and $\boldsymbol{B}_4$ have $\min(L_3, L_4) = 1$. Therefore we restrict further code searches to $\boldsymbol{B}_1$ and $\boldsymbol{B}_2$ cases only. The above optimization of feed-backward polynomials reduced the code search to a subset of $2^{10} = 1024$ candidate CCs for each of $\boldsymbol{B}_1$ and $\boldsymbol{B}_2$. The feed-forward polynomials $f_{2zk}(D)$, $z, k = 1, 2$ were further optimized in the next subsection applying design criteria from Subsection IV-A to the resultant DS of the STTuCM enumerated in Subsection III-B.

C. Distance Spectrum Optimization

To simplify the DS optimization, both CCs in the parallel concatenation were assumed to be the same. Also, in contrast to [15], [17] the N outputs of the second CC were altered during multiplexing, i.e., $c_n(l) = c_{N-n+1}^2(l)$ for l even [38]. In cases when error events in two CCs are both either odd- or even-punctured, such altering will, due to the systematic nature of CCs, assure that corners of the SCDM on the main diagonal have non-zero entries. This will presumably increase the number of candidate STTuCMs that satisfy the significant full rank criterion over quasi-static fading channels with no impact on the performance over fast fading channels.

To reduce complexity for the extensive code search over quasi-static fading channels, the cardinality of set $\mathcal{P}$ was reduced compared to the truncated UB analysis from Subsection III-C. To define $\mathcal{P}_{\text{quasi}}$ the following parameters were assumed: $h_{\max} = 3$, $\psi_{\max} = 8$, $\Delta_{\text{Pmax}} = 12$ and $\Theta_{\max} = 3$. Compared to Subsection III-C, $h_{\max}$ and Δ_{Pmax} were reduced, which was partly compensated by the slightly increased $\psi_{\max}$. For each of $\boldsymbol{B}_1$ and $\boldsymbol{B}_2$, 30 out of 1024 candidate CCs resulted in STTuCM that satisfied the significant full rank criterion. A further search over the minimum spectral product criterion was performed for each of the three EPD definitions from Subsection IV-A, i.e., assuming $\Delta_{\text{P}} = \det(\boldsymbol{A})$, $\Delta_{\text{P}} = \det(\boldsymbol{I} + \boldsymbol{A})$ or $\Delta_{\text{P}} = \text{trace}(\boldsymbol{A})$. Interestingly, for each of $\boldsymbol{B}_1$ and $\boldsymbol{B}_2$, the same two CCs satisfied the minimum spectral product criterion for all three EPD definitions. This concedes the equal eigenvalue criterion from [53]. Table I summarizes the results of a code search over quasi-static fading channels where $\boldsymbol{F} = [f_{212}(D); f_{211}(D); f_{222}(D); f_{221}(D)]$ includes polynomial coefficients of $f_{2zk}(D) = f_{2zk}^0 + f_{2zk}^1 D + \ldots + f_{2zk}^{\nu_k} D^{\nu_k}$. All four codes have the same value of Δ_{DS} for a given EPD definition.

On fast fading channels, the reduced cardinality set $\mathcal{P}_{\text{fast}}$ was determined by $h_{\max} = 4$, $\psi_{\max} = 8$, $\Delta_{\text{Pmax}} = 48$ and $\Theta_{\max} = 2$. Compared to the truncated UB from Subsection III-C, $h_{\max}$ and $\Theta_{\max}$ were decreased, which was partly compensated by increasing $\psi_{\max}$. For each of $\boldsymbol{B}_1$ and $\boldsymbol{B}_2$, the DS optimization produced two new CCs. The moderately high SNR=12dB was assumed to assure that optimization is performed in the error-floor region. Table II summarizes the results of a code search over fast fading channels. All four codes in Table II have the same value of the given performance measure.

TABLE I
THE BEST $\nu = 3$ QPSK RECURSIVE SYSTEMATIC CONSTITUENT CODES FOR 2 BPS/HZ STTUCM FROM OPTIMIZATION OVER QUASI-STATIC FADING CHANNELS.

CC	$\boldsymbol{B}$	$\boldsymbol{F}$	Full rank $\vert_{\mathcal{P}_{\text{quasi}}}$	$\Delta_{\text{DS}}\vert_{\det(\boldsymbol{A})}$	$\Delta_{\text{DS}}\vert_{\det(\boldsymbol{I}+\boldsymbol{A})}$	$\Delta_{\text{DS}}\vert_{\text{trace}(\boldsymbol{A})}$	$P_{\text{FER}}\vert_{\mathcal{P}_{\text{fast}}}$
Q_a	$[10;10;01;10]$	$[101;110;001;100]$	yes	2.5×10^{-3}	6.2×10^{-4}	5.4×10^{-2}	3.5×10^{-4}
Q_b	$[10;10;01;10]$	$[101;110;100;010]$	yes	2.5×10^{-3}	6.2×10^{-4}	5.4×10^{-2}	3.5×10^{-4}
Q_c	$[10;01;00;10]$	$[110;110;011;100]$	yes	2.5×10^{-3}	6.2×10^{-4}	5.4×10^{-2}	3.5×10^{-4}
Q_d	$[10;01;00;10]$	$[110;110;101;010]$	yes	2.5×10^{-3}	6.2×10^{-4}	5.4×10^{-2}	3.5×10^{-4}

TABLE II
THE BEST $\nu = 3$ QPSK RECURSIVE SYSTEMATIC CONSTITUENT CODES FOR 2 BPS/HZ STTUCM FROM OPTIMIZATION OVER FAST FADING CHANNELS.

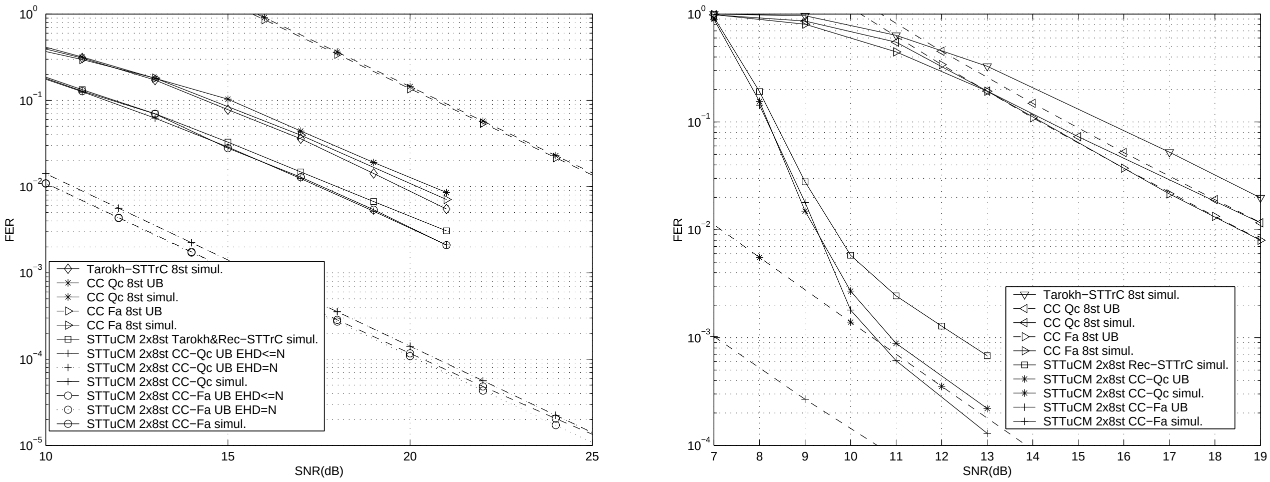
CC	$\boldsymbol{B}$	$\boldsymbol{F}$	Full rank $\vert_{\mathcal{P}_{\text{quasi}}}$	$\Delta_{\text{DS}}\vert_{\det(\boldsymbol{A})}$	$\Delta_{\text{DS}}\vert_{\det(\boldsymbol{I}+\boldsymbol{A})}$	$\Delta_{\text{DS}}\vert_{\text{trace}(\boldsymbol{A})}$	$P_{\text{FER}}\vert_{\mathcal{P}_{\text{fast}}}$
F_a	$[10;10;01;10]$	$[011;100;100;110]$	no	9.2×10^{-4}	3×10^{-4}	4.3×10^{-2}	2.5×10^{-5}
F_b	$[10;10;01;10]$	$[011;100;111;010]$	no	9.2×10^{-4}	3×10^{-4}	4.3×10^{-2}	2.5×10^{-5}
F_c	$[10;01;00;10]$	$[001;100;110;010]$	no	9.2×10^{-4}	3×10^{-4}	4.3×10^{-2}	2.5×10^{-5}
F_d	$[10;01;00;10]$	$[001;100;111;110]$	no	9.2×10^{-4}	3×10^{-4}	4.3×10^{-2}	2.5×10^{-5}

D. Performance evaluation

The newly designed CCs result in STTuCM with a 0.5 dB lower convergence threshold than with constituent Rec-STTrCs from [15], [17]. Fig. 6 depicts the FER performance of the STTuCM with the newly introduced CCs from Tables I and II over quasi-static and fast fading channels. As expected, different CCs from the same Table resulted in STTuCM with exactly the same performance. For brevity, only one curve per Table is plotted. The same parameters were adopted as in Subsection III-C. As seen from UB and simulation results in Fig. 6(a), new CCS from Table I resulted in full transmit diversity STTuCM with improved performance over both schemes studied in [15], [17]. New CCs from Table II offered additional gain over a fast fading channel and performed exactly the same as CCs from Table I over a quasi-static fading channel. Although STTuCM with CCs Table II did not satisfied the significant full rank criterion, the noticeable performance degradation was foreseen by UB only for large SNRs, high above those typically exhibited in practise. In summary, STTuCM with new CCs achieved more than 1.5 dB gain at FER 10^{-3} over both quasi-static and fast fading channels. The advantage of the bit-wise compared to symbol-wise interleaved STTuCM with the newly designed CCs over quasi-static and fast fading channels is shown in [38]. It was also demonstrated therein that by increasing the frame size, considerable coding gains, directly proportional to the number of independent fading realizations per frame, can be achieved for STTuCM. Replacing old with new CCs for STTuCM in [33], [34], [35] will inevitably lead to further performance improvement in future single- and multi-carrier based broadband wireless systems.

V. CONCLUSIONS

The performance of the STTuCM shown in simulations over a variety of fading channels motivated the theoretical studies in this paper. Both the fully recursive STTuCM with both CCs being Rec-STTrCs, and the non-fully recursive STTuCM where one of the CCs is replaced with the equivalent feed-forward STTrC, were studied. The performance analysis included the two-dimensional DS interpretation, the truncated UB and the iterative decoding convergence analysis. The DS and UB analysis discovered the effect of *spectral thinning* similar to single antenna turbo codes. This revealed the importance of multiplicities for the STTuCM design. It was further shown that without particular optimization of CCs and/or information interleaving, the punctured parallel concatenation in general does not preserve full transmit diversity over quasi-static fading channels. The superior performance of the non-fully recursive STTuCM over quasi-static fading channels was attributed to its less deteriorating rank deficient part of the DS and its better iterative decoding convergence properties. On fast fading channels, good two-dimensional DS properties were found to be crucial, accrediting the preferable fully-recursive STTuCM. The truncated UB appeared to be tight over fast fading channels but rather optimistic over quasi-static fading channels. Nevertheless, the rank deficient error events seem to have been well enumerated, which appeared to be crucial for the accurate design of the full transmit diversity STTuCM in the second half of the paper.



(a) (b)

Fig. 6. Performance of STTuCM with the new 8-state CCs, 10 decoding iterations, input information frame of $LK = 260$ bits, $N = 2$, $M = 1$. (a) Quasi-static fading channel. (b) Fast fading channel.

The modified design criteria for ST codes over quasi-static and fast fading channels were further introduced that capture the joint effects of error exponents and multiplicities in the code's DS. To assure the good convergence of iterative decoding, a binary recursive systematic form with a primitive equivalent feed-backward polynomial was assumed to model the subset of candidate CCs. The DS based optimization produced new sets of CCs that resulted in STTuCM with more than 1.5 dB gain at FER 10^{-3} over both quasi-static and fast fading channels. The full transmit diversity over quasi-static fading channels was achieved with no constrains on the implemented pseudo-random information interleaving. This leaves plenty of degrees of freedom for the eventual code-matched interleaving design.

REFERENCES

[1] E. Telatar, "Capacity of multi-antenna gaussian channels," *European Trans. Telecommun.*, vol. 10, no. 6, pp. 585–595, Nov.-December 1999.

[2] G. J. Foschini, "Layered space-time architecture for wireless communications in a fading environment when using multi-element antennas," *The Bell System Technical Journal*, pp. 41–59, 1996.

[3] H. Xu, M. Gans, D. Chiznik, J. Ling, P. Wolniansky, and R. Valenzuela, "Spatial and temporal variations of MIMO channels and impacts on capacity," in *Proc. IEEE Int. Conf. Commun.*, vol. 1, 2002, pp. 262–266.

[4] A. Molisch, M. Steinbauer, M. Toeltsch, E. Bonek, and R. S. Thom, "Capacity of MIMO systems based on measured wireless channels," *IEEE J. Select. Areas Commun.*, vol. 20, no. 3, pp. 561–569, Apr. 2002.

[5] D. Gesbert, H. Bolcskei, D. Gore, and A. Paulraj, "MIMO wireless channels: Capacity and performance prediction," in *Proc. IEEE Global Telecommun. Conf.*, vol. 2, 2000, pp. 1083 –1088.

[6] D. Chizhik, G. Foschini, M. Gans, and R. Valenzuela, "Keyholes, correlations, and capacities of multielement transmit and receive antennas," *IEEE Trans. Wireless Commun.*, vol. 1, no. 2, pp. 361–368, Apr. 2002.

[7] A. J. Paulraj and C. B. Papadias, "Space-time processing for wireless communications," *IEEE Signal Processing Mag.*, vol. 14, no. 6, pp. 49–83, Nov. 1997.

[8] A. F. Naguib, N. Seshadri, and A. R. Calderbank, "Increasing data rate over wireless channels," *IEEE Signal Processing Mag.*, vol. 17, no. 3, pp. 76–92, May 2000.

[9] D. Tujkovic, "Space-time turbo coded modulation for wireless communication systems," Ph.D. dissertation, University of Oulu, Oulu, Finland, 2003.

[10] V. Tarokh, N. Seshadri, and A. R. Calderbank, "Space-time codes for high data rate wireless communication: Performance criterion and code construction," *IEEE Trans. Inform. Theory*, vol. 44(2), no. 2, pp. 744–765, 1998.

[11] J. Grimm, M. P. Fitz, and J. V. Krogmeier, "Further results on space-time coding for Rayleigh fading," in *Proc. Annual Allerton Conf. Commun., Control, Computing*, Allerton House, Monticello, USA, Sept. 23-25 1998.

[12] S. H. Jamali and T. Lengoc, *Coded-modulation techniques for fading Channels.* Kluwer Academic Publishers, 1994.

[13] S. Benedetto, M. Mondin, and G. Montorsi, "Performance evaluation of trellis-coded modulation schemes," *Proc. IEEE*, vol. 82, no. 6, pp. 833–855, June 1994.

[14] B. Lu and X. Wang, "Space-time code design in OFDM systems," in *Proc. IEEE Global Telecommun. Conf.*, San Francisco, USA, Nov. 2000.

[15] D. Tujkovic, M. Juntti, and M. Latva-aho, "Space-time turbo coded modulation: Design and applications," *EURASIP Journal on Applied Signal Processing, Special Issue on Space-Time Coding and its Applications, Part I*, vol. 2002, no. 3, pp. 236–248, Mar. 2002.

[16] P. Robertson and T. Woerz, "Coded modulation scheme employing turbo codes," *Electron. Lett.*, vol. 31, pp. 1546–1547, Aug. 1995.

[17] D. Tujkovic, "Recursive space-time trellis codes for turbo coded modulation," in *Proc. IEEE Global Telecommun. Conf.*, vol. 2, San Francisco, USA, Nov. 27 - Dec. 1 2000, pp. 1010–1015.

[18] G. Battail, "A conceptual framework for understanding turbo codes," *IEEE J. Select. Areas Commun.*, vol. 16, no. 2, pp. 245–254, Feb. 1998.

[19] D. Tujkovic, "High bandwidth efficiency space-time turbo coded modulation," in *Proc. IEEE Int. Conf. Commun.*, vol. 4, Helsinki, Finland, June 11-14 2001, pp. 1104–1109.

[20] K. R. Narayanan, "Turbo decoding of concatenated space-time codes," in *Proc. Annual Allerton Conf. Commun., Control, Computing*, 1999.

[21] D. Cui and A. M. Haimovich, "Design and performance of turbo space-time coded modulation," in *Proc. IEEE Global Telecommun. Conf.*, 2000.

[22] W. Firmanto, Z. Chen, B. Vucetic, and J. Yuan, “Design of space-time turbo trellis coded modulation for fading channels,” in *Proc. IEEE Global Telecommun. Conf.*, 2001.
[23] D. Cui and A. M. Haimovich, “Error performance analysis of turbo space-time coded modulation over fading channels,” in *Proc. IEEE Int. Conf. Commun.*, 2001.
[24] ——, “Performance of parallel concatenated space-time codes,” *IEEE Commun. Lett.*, vol. 5, no. 6, pp. 236–238, June 2001.
[25] W. Firmanto, Z. Chen, and B. V. J. Yuan, “Space-time turbo trellis coded modulation for wireless data communications,” *EURASIP Journal on Applied Signal Processing, Special Issue on Space-Time Coding and its Applications, Part II*, vol. 2002, no. 5, pp. 459–470, May 2002.
[26] H. Ogiwara, A. Mizutome, and K. Koike, “Performance evaluation of parallel concatenated TCM,” in *Proc. Int. Symp. Turbo Cod. Related Topics*, 2000.
[27] T. Duman and M. Salehi, “Performance bounds for turbo-coded modulation systems,” *IEEE Trans. Commun.*, vol. 47, no. 4, pp. 511–521, Apr. 1999.
[28] ——, “The union bound for turbo-coded modulation systems over fading channels,” *IEEE Trans. Commun.*, vol. 47, no. 10, pp. 1495–1502, Oct. 1999.
[29] S. L. Goff, A. Glavieux, and C. Berrou, “Turbo-codes and high spectral efficiency modulation,” in *Proc. IEEE Int. Conf. Commun.*, vol. 2, 1994, pp. 645–649.
[30] A. Stefanov and T. Duman, “Turbo coded modulation for wireless communications with antenna diversity,” in *Proc. IEEE Veh. Technol. Conf.*, 1999, pp. 1565–1569.
[31] Y. Liu and M. P. Fitz, “Space-time turbo codes,” in *Proc. Annual Allerton Conf. Commun., Control, Computing*, 1999.
[32] H. J. Su and E. Geraniotis, “Space-time turbo codes with full antenna diversity,” *IEEE Trans. Commun.*, vol. 49, no. 1, pp. 47–57, Jan. 2001.
[33] D. Tujkovic, M. Juntti, and M. Latva-aho, “Space-frequency-time turbo coded modulation,” *IEEE Commun. Lett.*, vol. 5, no. 12, pp. 480–482, Dec. 2001.
[34] D. Tujkovic and E. Sottani, “Combined array processing and space-time turbo coded modulation for WCDMA downlink over frequency-selective Rayleigh fading channels,” in *Proc. IEEE Int. Symp. Spread Spectrum Techniques and Applications*, Prague, Czech Republic, Sept. 2002.
[35] ——, “Orthogonalized spatial multiplexing for WCDMA downlink over frequency-selective Rayleigh fading channels,” in *The 6th Baiona Workshop on Sign. Process. in Commun., submitted*, Baiona, Spain, Sept. 2003.
[36] L. Zexian, M. Vehkopera, D. Tujkovic, M. Latva-aho, and M. Juntti, “Performance evaluation of space-frequency coded MIMO MC-CDMA system,” in *Proc. IST Mobile Summit*, Aveiro, Portugal, to appear, June 2003.
[37] S. Siwamogsatham, M. P. Fitz, and J. H. Grimm, “A new view of performance analysis of transmit diversity schemes in correlated Rayleigh fading,” *IEEE Trans. Inform. Theory*, vol. 48, no. 4, pp. 950–956, Apr. 2002.
[38] D. Tujkovic, “Performance analysis and constituent code design for space-time turbo coded modulation over fading channels,” *IEEE Trans. Inform. Theory*, Mar. 2003, submitted. Available at `http://www.cwc.oulu.fi/%7edjordje/`.
[39] M. Simon, “Evaluation of average bit error probability for space-time coding based on a simpler exact evaluation of pairwise error probability,” *KICS J. Commun. Netw.*, vol. 3, no. 3, pp. 257–264, Sept. 2001.
[40] A. P. Rosiers and P. Siegel, “On performance bounds for space-time coded modulation on fading channels,” in *Proc. IEEE Int. Symp. Inform. Theory and its Applications*, Honolulu, Hawaii, USA, Nov.5–8 2000.
[41] D. Aktas and M. Fitz, “Computing the distance spectra of space-time trellis codes,” in *Proc. IEEE Wireless Commun. and Networking Conf.*, 2000, pp. 51–55.
[42] D. Kucukyavuz and M. Fitz, “New views of transfer function based performance analysis of coded modulations,” in *Proc. Conf. Asilomar Sign., Syst., Comp.*, 1999, pp. 42–46.
[43] G. Caire and E. Viterbo, “Upper bound on the frame error probability of terminated trellis codes,” *IEEE Commun. Lett.*, vol. 2, no. 1, pp. 2–4, Jan. 1998.
[44] S. Benedetto and G. Montorsi, “Unveiling turbo codes: Some results on parallel concatenated coding schemes,” *IEEE Trans. Inform. Theory*, vol. 42, no. 2, pp. 409–428, Mar. 1996.
[45] E. Malkamaki and H. Leib, “Evaluating the performance of convolutional codes over block fading channels,” *IEEE Trans. Inform. Theory*, vol. 45, no. 5, pp. 1643–1646, July 1999.
[46] L. Perez, J. Seghers, and D. J. Costello Jr., “A distance spectrum interpretation of turbo codes,” *IEEE Trans. Inform. Theory*, vol. 42, no. 6, pp. 1698–1709, Nov. 1996.
[47] D. Divsalar, S. Dolinar, and F. Polara, “Iterative turbo decoder analysis based on density evolution,” *IEEE J. Select. Areas Commun.*, vol. 19, no. 5, pp. 891–907, May 2001.
[48] H. E. Gamal and A. R. Hammons, “Analyzing the turbo decoder using Gaussian approximation,” *IEEE Trans. Commun.*, vol. 47, no. 2, pp. 2321–2334, Feb. 2001.
[49] Z. Chen, J. Yuan, and B. Vucetic, “Improved space-time trellis coded modulation scheme on slow Rayleigh fading channels,” *Electron. Lett.*, vol. 37, no. 7, pp. 440–441, Mar. 2001.
[50] M. Tao and R. S. Cheng, “Improved design criteria and new trellis codes for space-time coded modulation in slow flat fading channels,” *IEEE Commun. Lett.*, vol. 5, no. 7, pp. 313–315, July 2001.
[51] E. Biglieri, G. Taricco, and A. Tulino, “Performance of space-time codes for a large number of antennas,” *IEEE Trans. Inform. Theory*, vol. 48, no. 7, pp. 1794–1803, July 2002.
[52] D. M. Ionescu, “New results on space-time code design criteria,” in *Proc. IEEE Wireless Commun. and Networking Conf.*, 1999, pp. 684–687.
[53] D. M. Ionescu, K. Mukkavilli, Z. Yan, and J. Lilleberg, “Improved 8- and 16-state space-time codes for 4PSK with two transmit antennas,” *IEEE Commun. Lett.*, vol. 5, no. 7, pp. 301–303, July 2001.
[54] D. Tujkovic, “Constituent code optimization for space-time turbo coded modulation based on distance spectrum and iterative decoding convergence,” in *Proc. IEEE Inform. Theory Workshop*, vol. 1, Paris, France, Mar. - Apr. 2003, pp. 107–110.
[55] Y. S. Jung and J. H. Lee, “New measure of coding gain for space-time trellis codes,” in *Proc. IEEE Int. Symp. Inform. Theory*, Washington DC, USA, June 24–29 2001, p. 198.
[56] Y. Sasazaki and T. Ohtsuki, “Improved design criteria and new trellis codes on space-time trellis coded modulation in fast fading channels,” in *Proc. IEEE Global Telecommun. Conf.*, 2002.
[57] D. Tujkovic, “Constituent code optimization for space-time turbo coded modulation over fast fading channels,” in *Proc. IEEE Int. Symp. Inform. Theory*, Yokohama, Japan, June-July 2003.