Error Probability Bounds for Bit-Interleaved Space–Time Trellis Coding Over Block-Fading Channels

Yong Li and Jaekyun Moon, *Fellow, IEEE*

***Abstract*—In this correspondence, we investigate the error probability bounds of a bit-interleaved space–time trellis coding (BISTTC) scheme, which concatenates bit-interleaved coded modulation (BICM) with a space–time trellis code (STTC). We focus on general block-fading channels, wherein each data packet or frame spans a number of independent fading blocks. BICM applied to such channel environments can effectively exploit both time and frequency selectivity, while a STTC maximizes the spacial diversity order. The exact pairwise error probability (PEP) and weight enumeration function (WEF) of BISTTC are used to evaluate the error bound. Due to the concatenation of an outer error correction code (ECC) and a STTC, the overall WEF of BISTTC is obtained by combining the WEF of the outer code with that of the STTC through an uniform interleaver. The main challenge here is to compute the WEF of a STTC for block-fading channels with reasonable complexity. We rely on constructing a composite state transition matrix based on a number of single-step virtual trellises, each corresponding to an independent fading block within a frame. We discuss how this approach reduces storage and computational requirements in the bound analysis, compared to the existing method of obtaining the state transition matrix through accumulation of single-step transition matrices. The derived bound is applicable to both spatially uncorrelated and correlated channels as well as to both flat and frequency-selective block-fading channels. The bound is shown to provide a reasonably close estimate of the simulated performance based on the turbo equalizer-like iterative processing of soft information between the STTC decoder and the ECC decoder.**

***Index Terms*—Bit-interleaved coded modulation (BICM), bit-interleaved space–time trellis coding (BISTTC), block fading, weight enumeration function (WEF).**

I. Introduction

Diversity techniques are effective in combating multipath fading in wireless communication. Space–time coding schemes actively investigated in recent years are designed to provide diversity gains through two-dimensional coding over both the time and space dimensions [1]–[3]. In addition to maximizing the diversity order, the space–time trellis code (STTC) of [1] also provides coding gains. Space-time coding has also been specialized to orthogonal-frequency division-multiplexing (OFDM) applications, wherein coding is effectively applied in the frequency and space domains [4]–[6]. A somewhat different category of space–time methods also exists that is based on bit-interleaved coded modulation (BICM) [13] in conjunction with spatial-domain multiplexing (SM) of the transmitted symbols [14]. The examples along this approach include [8]–[12]. The original layered space-time method of [14] as well as its variations [15]–[17] also fall into this category in the sense that the layered space–time architectures typically assume outer coding and bit-level interleaving. The basic idea behind the BICM-based space–time methods is that spatial diversity is exploited indirectly through random distribution of information bits over different transmit antennas made possible by the interleaver operating on coded bits.

Manuscript received June 25, 2004; revised May 19, 2006.

Y. Li is at 5775 Morehouse Drive, San Diego, CA 92121 USA (e-mail: lixx0266@ece.umn.edu).

J. Moon is with the Department of Electrical and Computer Engineering, University of Minnesota, Minneapolis, MN 55455 USA (e-mail: moon@ece.umn.edu).

Communicated by B. Hassibi, Associate Editor for Communications.

Digital Object Identifier 10.1109/TIT.2007.907470

In [18], the present authors considered the concatenation of BICM with the existing STTC of [1]. The rationale for this scheme, dubbed bit interleaved space–time trellis coding (BISTTC), is that while the STTC ensures a maximum spatial diversity order, the presence of BICM will help further exploit both time diversity and frequency diversity (in the context of OFDM), when available. This scheme has practical merits in that it allows a flexible system architecture built upon known and available components. The comparative performance merits have also been discussed in [18].

The objective of this paper is to derive an error probability bound for such a BISTTC scheme over general block-fading channels, wherein each transmitted data frame contains a number of independently fading blocks. The error bound of the STTC itself has been investigated in the literature for fully interleaved (fast) fading channels [19], [20], quasistatic fading channels [21], [22], and some general cases including both types of fading channels [23]. Although the standard union bound is shown to be tight in fast fading channels, it tends to be very loose in quasistatic fading channels due to the lack of dominant error events in such channels [24]. Some methods have been suggested to alleviate this problem. In [25] an expurgated bound was proposed, although the expurgation alone is often insufficient to generate a tight bound in quasi-static channels. In [23] an approximated error bound was used that only considered error events with limited lengths. However, this truncated bound is no longer an upper bound, and convergences rather slowly to the actual union bound for quasistatic channels. Based on the idea originated from [24], another "limit before averaging" type technique was independently employed in [21] and [22], trying to tighten the union bound without truncation. This method results in much tighter union bounds, but requires large amounts of storage and computation when the block length is large and/or the associated STTC trellis is complex, since in these scenarios the number of error events to be examined grows large and the distance information for every error event with different distance needs be recorded.

To obtain error bounds for BISTTC, we need to find the weight enumeration function (WEF) of the outer code, compute the pairwise error probability (PEP) for error events, find the conditional WEF (CWEF) of the STTC, and then combine them appropriately. The combination is done through the notion of the uniform interleaver [26], which has been shown to be effective in characterizing the interleaver behavior in concatenated schemes. Since the WEF of the outer code such as the convolutional code (CC) is well established [26], the main challenge here is to handle the error events of the STTC in block-fading channels. In this paper, we also rely on the limitation approach, but introduce some new ideas that allow reduced computational and storage requirements in evaluating the bound for block-fading channels. Our approach starts with the definition of the error event metric based on Craig's formulation for the Q-function [30]. We then observe that for a group of error events corresponding to a given input weight but possibly having different metrics, it is the scalar "group metric" which is actually used in the bound evaluation. This group metric is a summation of the metrics of those error events belonging to that group weighted by their multiplicities. In this way the individual metrics as well as their multiplicities, which are the CWEF coefficients of the STTC, need not be explicitly enumerated; only the metric summation needs be recorded for subsequent processing. Therefore, the required storage and computational complexity are greatly reduced. In practice the metric for an error event can be decomposed into single-step state transition metrics which are associated with the state transition matrix of a STTC, and thus can be computed through a matrix accumulation process. For block-fading channels under investigation, we compute a composite state transition matrix to take into account the

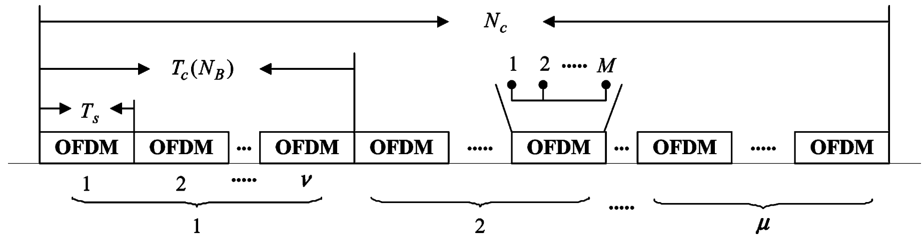

Fig. 1. Block fading channel model in the OFDM context.

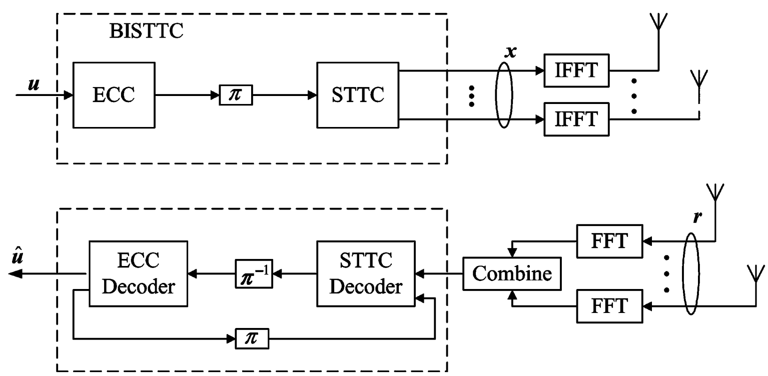

Fig. 2. System model of BISTTC.

time selectivity. The composite state transition matrix is obtained from a number of state transition matrices corresponding to the individual fading blocks within the frame. To obtain the state transition matrix for each fading block, a notion of the single-step virtual trellis is introduced. This idea was also mentioned in our previous work [18], but here the derived bounds are presented in detail as well as generalized to both flat and frequency-selective channels and to both spatially independent and correlated fading channels. As will be shown, the derived bounds provide good approximations to the error rate simulation results of the actual system employing a turbo-equalizer-like iterative demodulation/decoding scheme in above channel conditions. We note that the computational load of these bounds is still intense for complex STTC's achieving higher transmission rates. The difficulty of directly constructing such STTC's with a large constellation and/or a large number of states has also been recognized in [1], where the authors suggested a design of such STTC's through multi-level coding [27]. Of course, while computationally intensive, the bound still allows estimation of error probabilities in the high SNR region where a simulation study is not feasible.

The outline of this correspondence is as follows. In Section II, the system model of BISTTC is introduced. In Section III, the error bound is derived for block-fading channels. Simulation and numerical results are presented in Section IV. Finally, conclusions are drawn in Section V.

II. System Model

A. General Block Fading Channel Model

Here we adopt a general block-fading channel model of [28] that assumes that a given transmitted frame may span multiple independent fading blocks, each of which lasts for N_B symbols, as illustrated in Fig. 1. We use T_c and T_s to denote the channel coherence time and OFDM symbol duration, respectively. Also assume that a data frame spans μ independent fading blocks, during each of which the channel is static for ν OFDM symbol periods, i.e., $N_c = \mu N_B = \mu\nu M$, where M denotes the number of subcarriers in an OFDM symbol. In the following, we will use L to specify the number of taps in a tapped-delay-line channel model [29]. Note that when $L = 1$, the model corresponds to flat fading and subsequently $M = 1$ is implied since there exists no motivation for OFDM.

B. System Model

In Fig. 2, a BISTTC scheme is depicted with n_T transmit and m_R receive antennas. It is straightforward to see that BISTTC concatenates an outer error-correcting code (ECC) with an inner STTC, and at the receiver iterative decoding and demodulation (IDD) can be performed in a "turbo" fashion to further improve the performance.

The BISTTC codeword can be represented by 3-D indices: $x_{ip}^n(k)$, $1 \le k \le M$, $1 \le n \le n_T$, $1 \le i \le \mu$ and $1 \le p \le \nu$, denotes the symbol transmitted on the kth subcarrier from the nth transmit antenna during the pth OFDM symbol within the ith fading block. In the context of OFDM, the frequency response of the channel between transmit antenna n and receive antenna m, for the kth subcarrier of the pth OFDM symbol within the ith fading block, $H_{nm}(i,p,k)$, can be computed as

$$H_{nm}(i,p,k) = \sum_{l=1}^{L} g_{nm}(i,l) e^{-j2\pi(k-1)(l-1)/M}, \tag{1}$$

for $1 \le k \le M$, where $g_{nm}(i,l)$ denotes the lth-tap coefficient of the impulse response of the channel between the transmit antenna n and the receive antenna m during time block i, and they can be arranged in vector form as $\mathbf{g}_m^i = [\ldots, \mathbf{g}_m^{ilT}, \ldots]^T$ with $\mathbf{g}_m^{il} = [g_{1m}(i,l), \ldots, g_{n_T m}(i,l)]^T$, $1 \le l \le L$. Note that $H_{nm}(i,p,k)$'s are invariant with respect to p because in block-fading channels the time response $\mathbf{g}_m^i$'s are constant within the ith block.

III. Error Bound Performance Analysis

Every codeword of BISTTC is a sequence composed of STTC symbols, so the error events analysis of BISTTC basically follows that of the STTC. However, the concatenation with the outer code via an interleaver changes the weight distribution of the input sequence to the STTC encoder and manages to reduce the frequency at which the small-distance error events occur. Accordingly the weight enumeration function (WEF) of the outer code also need be taken into account. Due to the difficulty of characterizing a specific interleaver behavior in the analysis, the popular notion of the uniform interleaver [26] is employed, which has been shown to be very effective in analyzing the interleaver effect in concatenated schemes.

It is well known that the union bound on the frame error probability P_f may be calculated as [22]:

$$P_f \leq \int_{\mathbf{H}} \frac{1}{|\mathcal{S}|} \sum_{\mathbf{x}} \sum_{\hat{\mathbf{x}} \neq \mathbf{x}} P(\mathbf{x} \rightarrow \hat{\mathbf{x}}|\mathbf{H}) f(\mathbf{H}) \cdot d\mathbf{H} \tag{2}$$

where $|\mathcal{S}|$ denotes the number of BISTTC codewords, and $f(\mathbf{H})$ is the joint probability density function (pdf) of all channel coefficients contained in $\mathbf{H}$. Usually the Chernoff approximation to $P(\mathbf{x} \rightarrow \hat{\mathbf{x}}|\mathbf{H})$ is sufficient for PEP analysis and design purposes [1]; however, in order to tighten the error bound, an exact evaluation of the conditional PEP is employed here based on the Craig's formulation for $Q(x)$ [30], $Q(x) = \frac{1}{\pi}\int_0^{\pi/2} \exp(-x^2/2\sin^2\theta)d\theta$, as was done in [19]. Accordingly, (2) becomes

$$P_f \leq \int_{\mathbf{H}} \frac{1}{|\mathcal{S}|} \sum_{\mathbf{x}} \sum_{\hat{\mathbf{x}} \neq \mathbf{x}} \left[\frac{1}{\pi} \int_0^{\frac{\pi}{2}} \exp\left(-\frac{\Delta(\mathbf{x},\hat{\mathbf{x}})E_s}{4N_0 \sin^2\theta} \right) \cdot d\theta \right] f(\mathbf{H}) \cdot d\mathbf{H} \tag{3}$$

where the distance $\Delta(\mathbf{x},\hat{\mathbf{x}})$ is defined as

$$\Delta(\mathbf{x},\hat{\mathbf{x}}) = \sum_{i=1}^{\mu} \sum_{p=1}^{\nu} \sum_{k=1}^{M} \sum_{m=1}^{m_R} \left| \sum_{n=1}^{n_T} H_{nm}(i,p,k)\delta_{ipk}^n \right|^2 \tag{4}$$

with the symbol error $\delta_{ipk}^n = x_{ip}^n(k) - \hat{x}_{ip}^n(k)$.

In the standard union bound approach, (3) is evaluated by first performing averaging with respect to $\mathbf{H}$ before doing the summation. Assuming a Rayleigh distribution and following the standard diagonalization approach of [1], we would have

$$P_f \leq \frac{1}{\pi} \int_0^{\frac{\pi}{2}} \left[\frac{1}{|\mathcal{S}|} \times \sum_{\mathbf{x}} \sum_{\hat{\mathbf{x}} \neq \mathbf{x}} \prod_{\mathcal{R}(\mathbf{x},\hat{\mathbf{x}})} \left(1 + \lambda_{ii'} \frac{E_s}{4N_0 \sin^2\theta} \right)^{-m_R} \right] \cdot d\theta \tag{5}$$

where $\lambda_{ii'}$, $1 \leq i \leq \mu$, $1 \leq i' \leq n_T L$, denotes the eigenvalue of the "codeword difference matrix" for each pair $(\mathbf{x},\hat{\mathbf{x}})$ [18], and $\mathcal{R}(\mathbf{x},\hat{\mathbf{x}})$ is the set of $(\mathbf{x},\hat{\mathbf{x}})$ which will result in nonzero eigenvalues.

For fast fading channels, (5) can be computed based on the transfer function approach [19]. The union bound is tight in this case, but for general block-fading channels including the quasistatic channels, the union bound (5) would be very loose. Therefore, a limitation method can be applied to (3) as in [22], to obtain

$$P_f \leq \int_{\mathbf{H}} \min\left[1, \frac{1}{\pi} \int_0^{\frac{\pi}{2}} \frac{1}{|\mathcal{S}|} \times \sum_{\mathbf{x}} \sum_{\hat{\mathbf{x}} \neq \mathbf{x}} \exp\left(-\frac{\Delta(\mathbf{x},\hat{\mathbf{x}})E_s}{4N_0 \sin^2\theta} \right) d\theta \right] \cdot f(\mathbf{H}) d\mathbf{H}. \tag{6}$$

The basic idea behind the limitation in (6) is that the conditional union bound is only to be used when it gives reasonable results (such as less than 1); otherwise it is upper-bounded by 1. Due to the limitation, the order of the integration with respect to $\mathbf{H}$ and the summation in (6) can not be interchanged; a multidimensional numerical integration has to be carried out.

Furthermore, BISTTC under consideration in this paper employs existing STTC's of [1], which are geometrically uniform according to the Forney's criterion [31]. Thus the error bound can be evaluated assuming that an arbitrary codeword such as the all-zero codeword is transmitted, as was done in [19]. Then in evaluating (6) only the nonzero codewords instead of all pairs of codewords need to be considered. Denote $\Delta \equiv \Delta(\mathbf{0},\mathbf{x})$ as the Euclidean distance between codeword $\mathbf{x}$ and the all-zero codeword. Then the overall WEF of BISTTC is defined as

$$B^C(Z) = \sum_{\Delta} A_\Delta Z^\Delta \tag{7}$$

where the multiplicity A_Δ denotes the number of codewords with distance Δ. Similarly, the conditional WEF (CWEF) of BISTTC is defined as

$$B_w^C(Z) = \sum_{\Delta|w} A_{w,\Delta} Z^\Delta \tag{8}$$

where $A_{w,\Delta}$ denotes the number of codewords with distance Δ generated by an information bit sequence of weight w, and the set "$\Delta|w$" only picks up those Δ's corresponding to a particular w. In contrast to the way the WEF is traditionally defined, the WEF given here is based on the Euclidean distance instead of Hamming distance. The bound (6) is now rewritten as

$$P_f \leq \int_{\mathbf{H}} \min\left[1, \frac{1}{\pi} \int_0^{\frac{\pi}{2}} B^C(Z)|_{Z=z_0} d\theta \right] f(\mathbf{H}) d\mathbf{H} \tag{9}$$

where $z_0 = \exp\left(-\frac{E_s}{4N_0 \sin^2\theta} \right)$, and the error bound for BER as

$$P_b \leq \int_{\mathbf{H}} \min\left[1, \frac{1}{\pi} \int_0^{\frac{\pi}{2}} \sum_{w} \frac{w}{N} \cdot B_w^C(Z)|_{Z=z_0} d\theta \right] f(\mathbf{H}) d\mathbf{H} \tag{10}$$

where N is the frame length in number of information bits. Note that in flat-fading cases the variable of integration would change from $\mathbf{H}$ to $\mathbf{g}$.

To evaluate the bounds (9) and (10), we need to find the multiplicities A_Δ and $A_{w,\Delta}$, and then substitute $Z = z_0$ into $B^C(Z)$ and $B_w^C(Z)$ to get $B^C(z_0)$ and $B_w^C(z_0)$, respectively. Since BISTTC is a concatenated scheme, its WEF and CWEF can be conveniently obtained by combining the WEF and CWEF of the outer code with the CWEF of the STTC through an uniform interleaver [26], respectively. Let A_d^O and $A_{w,d}^O$ denote the number of outer codewords with Hamming weight d and the number of such codewords generated by an input information bit sequence of weight w, respectively. Similarly, let $A_{d,\Delta}^{ST}$ denote the number of STTC codewords with distance Δ generated by an input bit sequence of weight d, and $B_d^{ST}(Z) = \sum_{\Delta|d} A_{d,\Delta}^{ST} Z^\Delta$ be the CWEF of a STTC.

Next we take an example to illustrate how $B^C(z_0)$ is computed. Following [26], first we obtain the WEF of BISTTC as

$$\begin{aligned} B^C(Z) &= \sum_d A_d^O \cdot B_d^{ST}(Z) / \binom{N/R_o}{d} \\ &= \sum_{\Delta} \underbrace{\left(\sum_d A_d^O \cdot A_{d,\Delta}^{ST} / \binom{N/R_o}{d} \right)}_{A_\Delta} Z^\Delta \end{aligned} \tag{11}$$

where R_o is the code rate of the outer code, and then compute $B^C(z_0)$ by substituting $Z = z_0$ into $B^C(Z)$. Equation (11) clearly shows that both sets of coefficients A_d^O's and $A_{d,\Delta}^{ST}$'s need be explicitly enumerated to obtain the WEF of BISTTC.

We can alternatively compute $B^C(z_0)$ by

$$B^C(z_0) = \sum_d A_d^O \cdot B_d^{ST}(z_0) / \binom{N/R_o}{d} \tag{12}$$

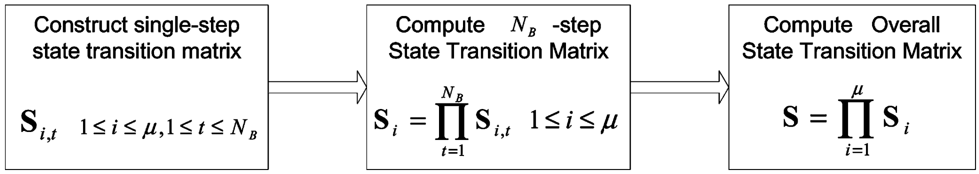

Fig. 3. Matrix accumulation process to compute $B_d^{ST}(z_0)$ for STTC.

where

$$B_d^{ST}(z_0) = \sum_{\Delta|d} A_{d,\Delta}^{ST} M_\Delta(\theta) \tag{13}$$

with a metric $M_\Delta(\theta) \triangleq \exp\left(-\frac{\Delta \cdot E_s}{4N_0 \sin^2\theta}\right)$. The approaches of (11) and (12) are conceptually equivalent, but in actual computation, the method of (12) directly calculates $B_d^{ST}(z_0)$ as a scalar summation through the state transition diagram, and thus coefficients $A_{d,\Delta}^{ST}$ need not be explicitly enumerated. Similarly, $B_w^C(z_0)$ can be computed as

$$B_w^C(z_0) = \sum_d A_{w,d}^O \cdot B_d^{ST}(z_0) / \binom{N/R_o}{d}. \tag{14}$$

This method results in a great reduction of required storage and computational complexity, as will be elaborated in later sections.

Assuming an outer convolutional code (CC), its WEF and CWEF can be conveniently obtained from its state diagram by an effective algorithm discussed in [26]. Thus the key issue in evaluating (12) and (14) is to find $B_d^{ST}(z_0)$, which is a weighted summation of the metrics $M_\Delta(\theta)$'s for certain error events. Next we shall focus on the decomposition of $M_\Delta(\theta)$ into a product form, and how to compute $B_d^{ST}(z_0)$.

A. Decomposition of $M_\Delta(\theta)$

Following the approach in [1], the distance Δ of (4) corresponding to a particular error event can be computed as

$$\Delta = \sum_{i=1}^{\mu} \sum_{t=1}^{N_B} \left(\sum_{m=1}^{m_R} \mathbf{H}_m^{it\,\mathcal{H}} \mathbf{D}_{it} \mathbf{H}_m^{it} \right) \tag{15}$$

where $\mathbf{H}_m^{it}$ is a vector containing those $H_{nm}(i,p,k)$'s in (4), t is the virtual-time index [6]: $t = (p-1)M + k, 1 \le p \le \nu, 1 \le k \le M$, and $\mathbf{D}_{it}$ is the codeword difference matrix during the ith block corresponding to the index t.

Substituting (15) into $M_\Delta(\theta)$ in (13), we will have

$$M_\Delta(\theta) = \prod_{i=1}^{\mu} \prod_{t=1}^{N_B} \prod_{m=1}^{m_R} \exp\left(-\frac{E_s \cdot \mathbf{H}_m^{it\,\mathcal{H}} \mathbf{D}_{it} \mathbf{H}_m^{it}}{4N_0 \sin^2\theta} \right). \tag{16}$$

Note that in the simpler case of flat-fading channels, $\mathbf{H}_m^{it} = \mathbf{g}_m^i$, is fixed within the ith fading block, so (16) becomes

$$M_\Delta(\theta) = \prod_{i=1}^{\mu} \prod_{t=1}^{N_B} \prod_{m=1}^{m_R} \exp\left(-\frac{E_s \cdot \mathbf{g}_m^{i\;\mathcal{H}} \mathbf{D}_{it} \mathbf{g}_m^i}{4N_0 \sin^2\theta} \right), \tag{17}$$

and a complete derivation following that for flat-fading channels has been presented in [18].

B. Steps to Compute $B_d^{ST}(z_0)$ for a STTC

A straightforward way to evaluate $B_d^{ST}(z_0)$ is, for a given θ, to compute $M_\Delta(\theta)$ for each error event and count its multiplicity. However, in practice the exhaustive search for all error events need not be performed; instead $B_d^{ST}(z_0)$ can be evaluated through the state diagram of a STTC. The explicit expression of $M_\Delta(\theta)$ in a product form makes it conveniently associated with the state transition matrix, which can be obtained from the state diagram of a STTC. In [22] the metric is obtained in a similar manner but from the *extended* state diagram, which enumerates every pair of error events under the assumption that any codeword can be transmitted. However, since the STTC considered here is geometrically uniform, an equivalent approach based on the state diagram can be used, which only enumerates nonzero error events under the premise that the all-zero codeword is transmitted [19], resulting in much lower computational complexity. We shall illustrate in the following how the $B_d^{ST}(z_0)$ of a STTC over general frequency-selective block-fading channels can be obtained. The result for flat-fading channels will also be mentioned as a special case.

The evaluation of $B_d^{ST}(z_0)$ via the state diagram of a STTC is to accumulate the state transition matrix step by step, which is basically a reverse process of decomposing $M_\Delta(\theta)$. The multiplicities associated with $M_\Delta(\theta)$ will be automatically captured during the process and are not counted explicitly in the process. For generalized block-fading channels, this evaluation process can be described by a flow chart in Fig. 3, and computation details can be found as follows.

1) *Step 1:* Consider a typical STTC scheme in [1] with 2^v states. The single-step state transition matrix of the STTC for the slot $\langle i,t \rangle$, $1 \le i \le \mu$, $1 \le t \le N_B$ is given by $\mathbf{S}_{it}(I,\theta) \triangleq \left[s_{it}^{ab}(I,\theta) \right]$, where I is a dummy variable whose exponent will serve to collect the Hamming weight of the input bits to the STTC, and the entry at row a and column b, $s_{it}^{ab}(I,\theta)$, $a,b \in \{1,2,\ldots,2^v\}$, indicates the single-step transition from the state s_a to the state s_b during slot $\langle i,t \rangle$ and is labeled by

$$s_{it}^{ab}(I,\theta) = \begin{cases} I^{d^{ab}} \cdot M_{\Delta_{it}^{ab}}(\theta), & \text{if } s_a \to s_b \text{ exists} \\ 0, & \text{otherwise} \end{cases} \tag{18}$$

where $M_{\Delta_{it}^{ab}}(\theta)$ and d^{ab} denote the distance metric and input Hamming weight associated with the transition $s_a \to s_b$, respectively. For a given θ the $M_{\Delta_{it}^{ab}}(\theta)$ is a *scalar*. In fact $\mathbf{D}_{it}^{ab}$ does not change with t, but in frequency-selective fading channels, $\mathbf{H}_m^{it}$ varies with t, and thus $s_{it}^{ab}(I,\theta)$ still need be computed individually for each t.

2) *Step 2:* Next we compute the composite N_B-step state transition matrix of the STTC for the ith block, $\mathbf{S}_i(I,\theta)$. It is easy to see that for a given θ, the entry at row a and column b of $\mathbf{S}_i(I,\theta)$, $s_i^{ab}(I,\theta)$, is a polynomial in I whose exponents are obtained by collecting the exponents of certain I-terms in $\mathbf{S}_{it}(I,\theta)$'s entries, and whose coefficients are the accumulated scalar multiplication of the coefficients of the corresponding I-terms. Conceptually, $s_i^{ab}(I,\theta)$ represents an aggregation of some "partial" error events, each of which starts from s_a and ends at s_b within the ith block. Consequently, the N_B-step trellis of the ith block is condensed into a one-step "virtual trellis" with the branch metric of transition $s_a \to s_b$ labeled as $s_i^{ab}(I,\theta)$. See Fig. 4.

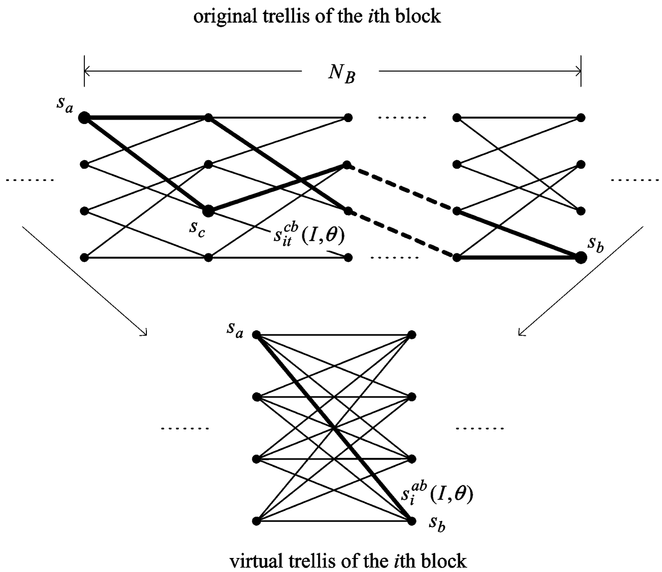

Fig. 4. Virtual trellis of STTC for block-fading channels.

To be more specific, the aggregated entry $s_i^{ab}(I,\theta)$ can be expressed as

$$s_i^{ab}(I,\theta)=\sum_{d_i^{ab}} I^{d_i^{ab}}\cdot M_{d_i^{ab}}(\theta) \tag{19}$$

where

$$M_{d_i^{ab}}(\theta)=\sum_{\Delta_i^{ab}|d_i^{ab}} A^{ST}_{d_i^{ab},\Delta_i^{ab}} M_{\Delta_i^{ab}}(\theta). \tag{20}$$

For a given θ, every $M_{d_i^{ab}}(\theta)$ is still a scalar. In actual computation $A^{ST}_{d_i^{ab},\Delta_i^{ab}}$ and $M_{\Delta_i^{ab}}(\theta)$ need not be explicitly enumerated; the coefficient associated with $M_{d_i^{ab}}(\theta)$ will be automatically obtained when the polynomial multiplication is done. Therefore, for a group of error events having a particular d_i^{ab}, we do not have to store the distance information Δ_i^{ab}'s and the multiplicities $A^{ST}_{d_i^{ab},\Delta_i^{ab}}$'s individually as in [22]; instead we only need store the scalar metric summation $M_{d_i^{ab}}(\theta)$ for that group.
This approach has two advantages. First, the reduction of storage requirement is obvious. Second, as will be shown next, the state transition matrices $\mathbf{S}_i(I,\theta)$'s of different fading blocks need be further multiplied to obtain the overall state transition matrix of the whole frame. So a reduced number of terms to be summed in each entry of $\mathbf{S}_i(I,\theta)$ will result in a reduced number of multiplications needed in the subsequent computation. For a STTC with a large number of states and/or a channel with a large block length, the reduction in storage and computational complexity of this algorithm would be large.

3) *Step 3:* Now we compute the overall state transition matrix of the STTC for the whole frame, i.e., $\mathbf{S}(I,\theta)$. The entry at row a and column b of $\mathbf{S}(I,\theta)$, $s^{ab}(I,\theta)$, will carry the information on those "complete" error events starting from s_a and ending at s_b within the given frame. In practice, all error events both start from and terminate to the zero state, and thus we only need keep the $(1,1)$ entry of $\mathbf{S}(I,\theta)$. Dropping the superscript indicating the state transition for the moment, i.e., $s(I,\theta)\equiv s^{11}(I,\theta)$, we write

$$s(I,\theta)=\sum_d I^d\cdot M_d(\theta) \tag{21}$$

where

$$M_d(\theta)=\sum_{\Delta|d} A^{ST}_{d,\Delta} M_\Delta(\theta). \tag{22}$$

As explained before, this entry is an enumeration by groups of all error events starting and ending at the zero-state, with each group sharing a particular input weight d, and each "group metric" being the summation of the metrics for those error events belonging to that group. Arranging them according to d, the coefficients of I^d, $M_d(\theta)$'s, will be the sets of $B_d^{ST}(z_0)$ in (13). Then we can combine these $B_d^{ST}(z_0)$'s with the WEF and CWEF of the outer code as in (12) and (14), and thus obtain the $B^C(z_0)$ and $B_w^C(z_0)$ of BISTTC to evaluate the error bounds.

C. Numerical Considerations

1) Integration: To evaluate the error bounds, the multi-dimensional integration with respect to $\mathbf{H}$ is done numerically. Although there are efficient routines available, generally we prefer the Monte Carlo method due to its simplicity, especially when the number of integration variables grows. In this way, the variables of integration contained in $\mathbf{H}$ or even $\mathbf{g}$ need not be independent, as we do not intend to decompose the multidimensional integral into a product of one-dimensional integrals. So the bounds derived here are naturally generalized to be applicable to both uncorrelated and correlated fading channels.

2) Expurgation: The expurgated bound [25] is usually used, since it only counts the *simple* error events diverging from and reemerging to the zero state for only once so that results in computation reduction, while still serving as an upper bound [22]. However, here we do not use the expurgation approach for the following reasons. First of all, in a concatenated system like in [26], there is no guarantee that simple error events in the inner trellis will correspond to likely error events in a global sense. Furthermore, even for the inner code only, due to block fading, we simply have no means to track the frame-wise simple error events. This is because that the error events for the whole frame are not directly enumerated; each frame-wise error event is rather made up of some block-wise error events concatenated to one another.

3) Truncation: In practice, counting all error events is unnecessary. With the limitation method applied, the error bound based on a few error events converges to a stable value. With truncation and limitation, this particular bound may not be conceptually considered as an upper bound, however, in practice it will serve as a reasonably good performance estimate of the simulated system. Theoretical explanation of this convergence behavior is under investigation.

IV. Simulation and Numerical Results

The BISTTC example investigated in this paper is the concatenation of an outer CC $(7,5)$ with a four-state QPSK STTC of [1]. We choose $n_T=2, m_R=1$, and an interleaver size of 1024 throughout the investigation. The figure of merit here is based on the block error probability (BLEP), defined as the normalized frame error rate (FER): $\mathrm{BLEP}=\mathrm{FER}/\mu$. The FER is defined as the probability that at least one bit error occurs within a frame. Due to the normalization, BLEP becomes independent of the parameter μ, allowing direct and reasonable performance comparison among situations with different μ values. The signal-to-noise ratio (SNR) is defined as $\mathrm{SNR}=n_T E_s/N_0$.

Both spatially uncorrelated and correlated channels are considered. Also, both frequency-flat and frequency-selective fading channels are investigated. In the frequency-selective fading case, OFDM with $M=64$ is assumed. A multiray equal-power channel model is used, in which each of the L channel taps is an i.i.d complex Gaussian random variable with zero-mean and variance $2\sigma_l^2=1/L$. We assume that the channel response is invariant for at least one OFDM symbol period.

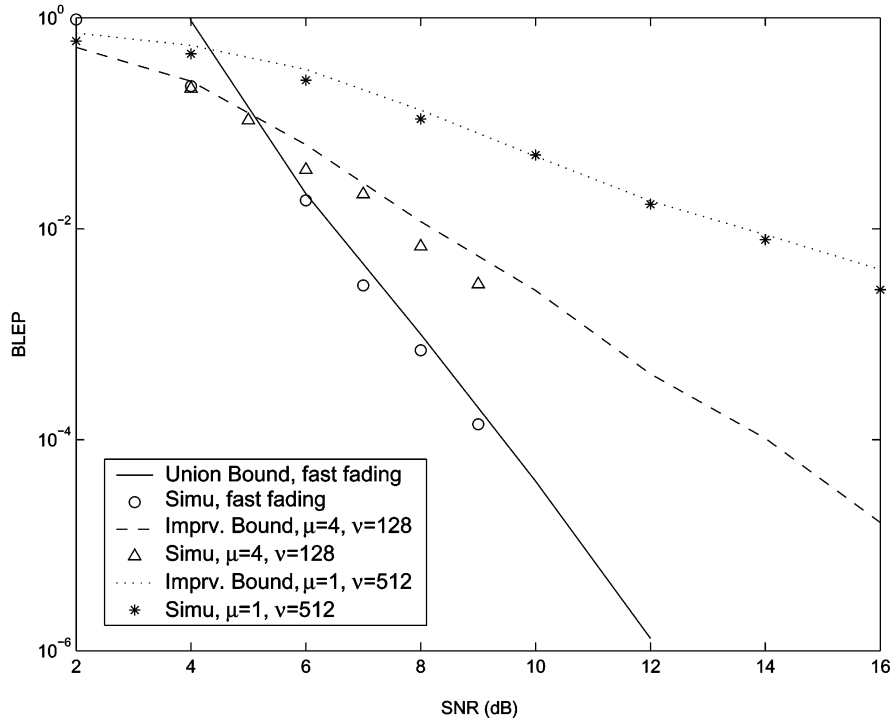

Fig. 5. Bound and simulation of BISTTC over spatially uncorrelated flat-fading channels, $L = M = 1$.

Note that the simulated performance of BISTTC is based on IDD scheme, which is a suboptimal algorithm but may approach the maximum-likelihood (ML) performance with sufficient iterations (five iteration in this paper). For comparison, the error bound derived in Section III is based on ML decoding.

A. Spatially Uncorrelated Frequency Flat-Fading Channels

The results are shown in Fig. 5. For fast fading, only the standard union bound is shown since it is indistinguishable from the improved bound of (9). Although not shown, for block-fading channels with a short block length, i.e., $N_B < 10$ symbols, the union bound still provides good estimate for the error probability. However, for block-fading channels with more realistic block lengths, such as $\mu = 4$, $\nu = 128$ and $\mu = 1$, $\nu = 512$ (quasistatic), the standard union bounds diverge. On the other hand, as shown in Fig. 5, the derived upper bound approximates the actual performance (based on IDD) accurately.

B. Spatially Uncorrelated Frequency Selective-Fading Channels

We consider a multipath-fading channel with $L = 4$. OFDM is used to remove the intersymbol interference (ISI) and convert the channel into parallel flat-fading channels. Note that in this case the variables of integration $\mathbf{H}$ in (9) are correlated for different subcarriers of an OFDM system, but still assumed to be spatially uncorrelated among transmitter antennas. The performances with different μ are shown in Fig. 6, where the performance curve of $L = 1, \mu = 1$ is also shown as a reference. It can be seen that in the frequency-selective fading channel, the derived bound provides an estimate of the IDD-based actual error rate with the accuracy of within approximately 1 dB of SNR, for both quasistatic $(\mu = 1)$ and block $(\mu = 8)$ fading channels. The plots clearly indicate that when more time or frequency diversity is available the error rate performance is improved [18].

C. Spatially Correlated Fading Channels

Let us now consider a more general case, where the channels could be correlated between different transmit–receive antenna pairs. Examples of BISTTC on both the flat fading $(L = 1)$ and frequency-selective fading $(L = 4)$, quasistatic $(\mu = 1)$ and block fading $(\mu = 4, 8)$ channels are considered. The correlation factor r is specified by

$$E\left\{\mathbf{g}_1^{il}\mathbf{g}_1^{il\mathcal{H}}\right\} = \begin{bmatrix} 1 & r \\ r & 1 \end{bmatrix} \text{ for } 1 \le i \le \mu, \quad 1 \le l \le L \tag{23}$$

where $\mathbf{g}_1^{il} = [g_{11}(i,l), g_{21}(i,l)]^T$ is defined in Section II-B. Fig. 7 shows the result for $r = 0.85$, which indicates that while the accuracy in estimating the error probability is reduced, the derived bounds again reflect the general trend for the error probability quite well. The enlarged gap between the simulated performance and ML upper bound is possibly due to the fact that IDD degrades more in correlated channels. Both the bound and the simulated curves show that the performance is improved with increased time and frequency diversity orders, as expected. The error rate performance of a STTC on such spatially correlated flat-fading channels has also been observed in [22].

D. Discussion

As indicated in Section III-C, although error events with length up to N_B need be examined in the algorithm for obtaining the WEF of STTC, we find from our simulation that it is usually sufficient to only enumerate those error events with length up to 30 symbols. The value might be higher when going for a more complex STTC, but the bounds will still converge possibly due to the use of limitation method.

BISTTC examined in this paper is composed of a simple CC and 4-state QPSK-based STTC. The bound for BISTTC based on more complex STTC, such as 16-state 16QAM-based STTC, can also be evaluated. However, beyond that, the computational load grows and may get out of hand, since the number of error events need be examined becomes too high even with the error event truncation. The main bottleneck is the number of states and the constellation size, which determine the dimension and the number of nonzero entries of the state transition matrix for STTC, respectively. Note that n_T can be extended to larger than 2, and the computational load of obtaining the bound will

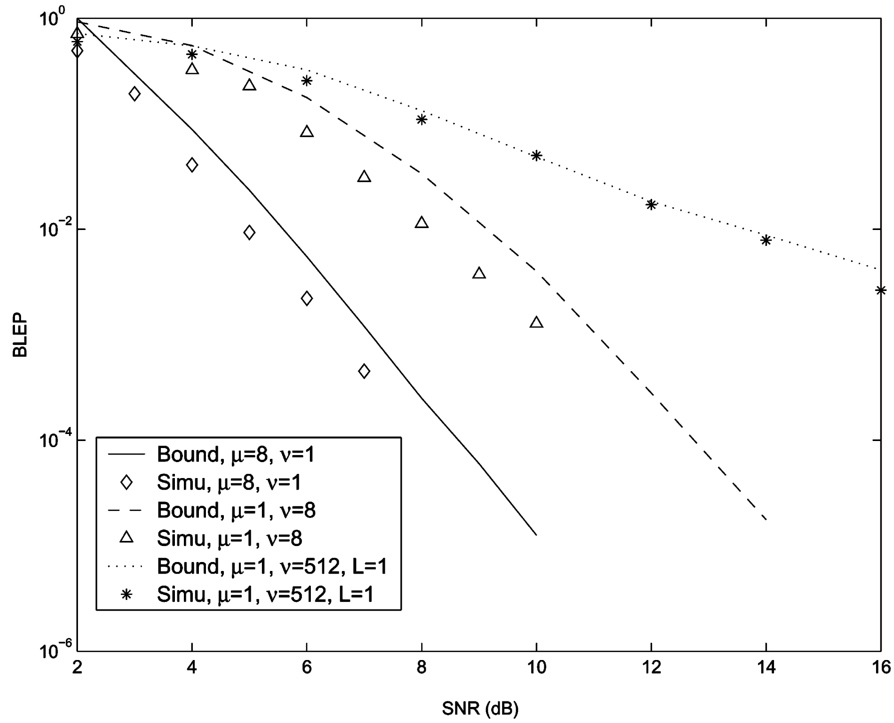

Fig. 6. Bound and simulation of BISTTC over spatially uncorrelated frequency-selective fading channels, $L = 4, M = 64$.

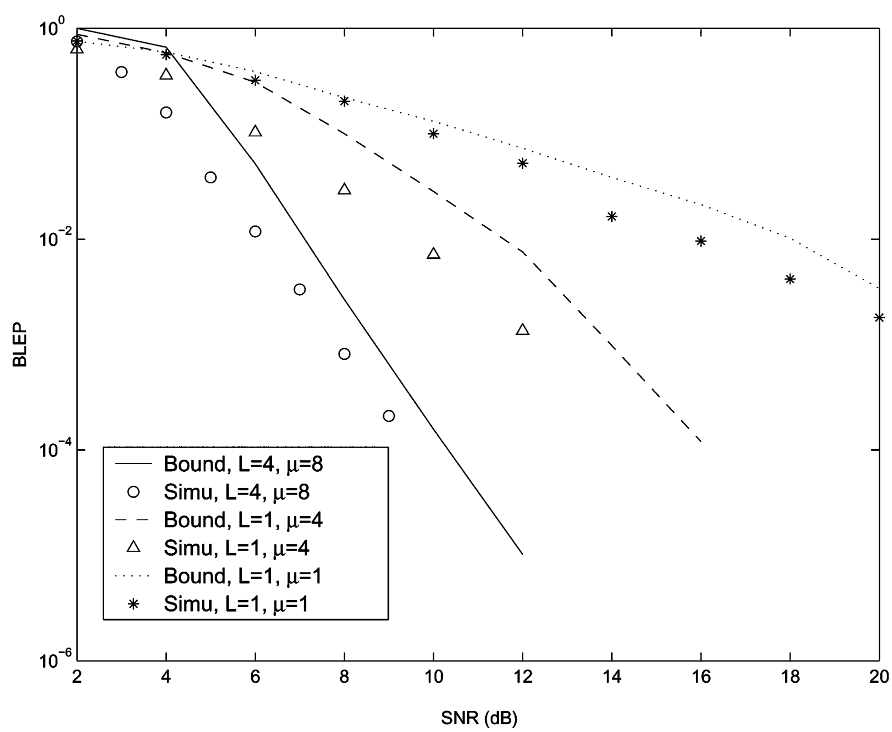

Fig. 7. Bound and simulation of BISTTC over spatially correlated fading channels with $r = 0.85$.

not increase since the trellis complexity of STTC does not change by merely increasing n_T.

On the other hand, it should be noted that more complex STTCs beyond 16-state 16QAM are rarely used in real systems, since the design becomes highly impractical in any case. For STTCs in practical use, the error bound derived here provides an efficient means for performance analysis, especially for the error rate regions where simulation becomes infeasible.

V. Conclusion

We have conducted the error bound analysis of an OFDM-based BISTTC scheme which concatenates BICM with known STTCs, over the generalized block-fading channels, wherein a single uninterrupted transmission of data stream encounters a number of independent block fading modes. A limitation method has been utilized to prevent the standard union bound from diverging. To evaluate the bounds, the exact PEP and the overall WEF of BISTTC have been used. The WEF of BISTTC can be obtained by combing the WEF of the outer code and the CWEF of STTC through an uniform interleaver. A reduced-complexity algorithm is derived to obtain the CWEF of the STTC through its single-step state transition matrix. The idea is based on constructing a one-step virtual trellis for each fading block in the frame, from which a block-specific state transition matrix is created that will in turn be used to compute the overall state transition matrix for the frame. The derived bound is shown to be able to reasonably capture the simulated performance of BISTTC, in both spatially uncorrelated and correlated as well as flat- and frequency-selective block-fading channels.

REFERENCES

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

[2] V. Tarokh, H. Jafarkhani, and A. R. Calderbank, "Space-time block codes from orthogonal designs," *IEEE Trans. Inf. Theory*, vol. 45, no. 5, pp. 1456–1467, Jul. 1999.

[3] S. M. Alamouti, "A simple transmit diversity technique for wireless communications," *IEEE J. Sel. Areas Commun.*, vol. 16, no. 8, pp. 1451–1458, Oct. 1998.

[4] H. Bolcskei and A. Paulraj, "Space-frequency coded broadband OFDM systems," in *Wireless Commun. and Networking Conf.*, Chicago, IL, Sep. 2000, pp. 1–6.

[5] B. Lu and X. Wang, "Space-time code design in OFDM systems," in *Proc. GLOBLECOM*, San Francisco, CA, Nov. 2000, vol. 2, pp. 1000–1004.

[6] Z. Liu, Y. Xin, and G. B. Giannakis, "Space-time-frequency coded OFDM over frequency-selective fading channels," *IEEE Trans. Signal Processing*, vol. 50, no. 10, pp. 2465–2476, Oct. 2002.

[7] J. A. C. Bingham, "Multicarrier modulation for data transmission: An idea whose time has come," *IEEE Commun.. Mag.*, pp. 5–14, May 1990.

[8] E. Biglieri, G. Taricco, and E. Viterbo, "Bit-interleaved time-space codes for fading channels," in *Proc. CISS*, Mar. 2000, pp. WA4.1–4.6.

[9] A. M. Tonello, "Space-time bit-interleaved coded modulation with an iterative decoding strategy," in *Proc.IEEE Veh.Technol. Conf.*, Boston, Sep. 2000, pp. 473–478.

[10] D. Park and B. G. Lee, "Design criteria and performance of space-frequency bit-interleaved coded modulations in frequency-selective Rayleigh fading channels," *J.Commun. Netw.*, vol. 5, no. 2, pp. 141–149, Jun. 2003.

[11] 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. Sel. Areas Commun.*, vol. 19, no. 5, pp. 958–968, May 2001.

[12] B. Lu, X. Wang, and K. R. Narayanan, "LDPC-based space-time coded OFDM systems over correlated fading channels: Performance analysis and receiver design," *IEEE Trans. Commun.*, vol. 50, no. 1, pp. 74–88, Jan. 2002.

[13] G. Caire, G. Taricco, and E. Biglieri, "Bit-interleaved coded modulation," *IEEE Trans. Inf. Theory*, vol. 44, no. 3, pp. 927–946, May 1998.

[14] G. J. Foschini, "Layered space-time architecture for wireless communication in a fading environment when using multi-element antennas," *Bell Labs Tech. J.*, pp. 41–59, 1996.

[15] H. E. Gamal and A. R. Hammons, Jr, "A new approach to layered space-time coding and signal processing," *IEEE Trans. Inf. Theory*, vol. 47, pp. 2321–2334, Sep. 2001.

[16] Z. Liu and G. B. Giannakis, "Layered space-time coding design for high data rate transmissions," in *Proc. MILCOM*, McLean, VA, Oct. 28–31, , pp. 1295–1299.

[17] L. H.-J. Lampe, R. Schober, and R. F. H. Fischer, "Multilevel coding for multiple-antenna transmission," *IEEE Trans. Wireless Commun.*, vol. 3, no. 1, pp. 203–208, Jan. 2004.

[18] Y. Li and J. Moon, "Performance analysis of bit-interleaved space-time coding for OFDM in block fading channels," in *IEEE VTC04 Spring*, Milan, Italy, May 2004.

[19] H. Shin and J. H. Lee, "Upper bound on the error probability for space-time codes in fast fading channels," in *Proc. IEEE VTC*, Vancouver, Canada, Sep. 2002, pp. 243–246.

[20] M. Uysal and C. N. Georghiades, "Error performance analysis of space-time codes over rayleigh fading channels," in *IEEE VTC-Fall*, Sep. 2000, vol. 5, pp. 2285–2290.

[21] A. P. des Rosiers and P. H. Siegel, "Space-time code performance bounds on quasistatic channels," in *ICC*, Anchorage, AK, May 2003, pp. 3160–3164.

[22] A. Stefanov and T. M. Duman, "Performance bounds for space-time trellis codes," *IEEE Trans. Inf. Theory*, vol. 49, no. 9, pp. 2134–2140, Sep. 2003.

[23] M. Uysal and C. N. Georghiades, "On the error performance analysis of STTC: An analytical framework," in *Proc. IEEE WCNC*, Orlando, FL, Mar. 2002, vol. 1, pp. 99–104.

[24] E. Malkamaki and H. Leib, "Evaluating the performance of convolutional codes over block fading channels," *IEEE Trans. Inf. Theory*, vol. 45, pp. 1643–1646, Jul. 1999.

[25] D. Aktas and M. P. Fitz, "Computing the distance spectrum of space-time trellis codes," in *Proc. IEEE WCNC*, Chicago, IL, Sep. 2000, pp. 51–55.

[26] S. Benedetto and G. Montorsi, "Unveiling urbo codes: Some results on parallel concatenated coding schemes," *IEEE Trans. Inf. Theory*, vol. 42, no. 2, pp. 409–428, Mar. 1996.

[27] A. Imai and S. Hirakawa, "A new multilevel coding method using error control codes," *IEEE Trans. Inf. Theory*, vol. 23, pp. 371–377, May 1977.

[28] E. Biglieri, G. Caire, and G. Taricco, "Limiting performance of block-fading channels with multiple antenna," *IEEE Trans. Inf. Theory*, vol. 47, no. 4, pp. 1273–1289, May 2001.

[29] J. G. Proakis, *Digital Communications*. Singapore: McGraw-Hill, 1995.

[30] M. K. Simon and D. Divsalar, "Some new twists to problems involving the Gaussian probability integral," *IEEE Trans. Commun.*, vol. 46, pp. 200–210, Feb. 1998.

[31] G. D. Forney, Jr, "Geometrically uniform codes," *IEEE Trans. Inf. Theory*, vol. 37, pp. 1241–1260, Sep. 1991.