Correspondence

Space-Time Bit-Interleaved Coded Modulation for OFDM Systems

Inkyu Lee, Albert M. Chan, and Carl-Erik W. Sundberg

***Abstract*—Space-time coding techniques significantly improve transmission efficiency in radio channels by using multiple transmit and/or receive antennas and coordination of the signaling over these antennas. Bit-interleaved coded modulation gives good diversity gains with higher order modulation schemes using well-known binary convolutional codes on a single transmit and receive antenna link. By using orthogonal frequency division multiplexing (OFDM), wideband transmission can be achieved over frequency-selective fading radio channels without adaptive equalizers. In this correspondence, we combine these three ideas into a family of flexible space-time coding methods. The pairwise error probability is analyzed based on the correlated fading assumption. Near-optimum iterative decoders are evaluated by means of simulations for slowly varying wireless channels. Theoretical evaluation of the achievable degree of diversity is also presented. Significant performance gains over the wireless local area network (LAN) 802.11a standard system are reported.**

***Index Terms*—Iterative decoding, OFDM, space-time codes.**

I. Introduction

Wireless channels exhibit a number of severe impairments, among which, fading is one of the most severe. In addition, additive noise and interference have to be combatted. For narrowband channels, the fading can often be assumed to be flat, while for wideband channels, the fading is typically frequency selective. Much progress has been made to combat these types of impairments. Diversity is a classic method to improve transmission over fading channels [1]. More recently, so-called space-time coding methods [2]–[6] have been proposed to obtain both space and time diversity by using multiple receive and transmit antennas combined with matching modulation and coding.

Coded modulation schemes are methods that, e.g., utilize time diversity on single transmit and receive antenna radio links [7], [8]. There are ways to efficiently utilize binary convolutional codes to obtain diversity gains with higher order nonbinary modulation symbols. Among these, we point to bit-interleaved coded modulation (BICM) systems [9], [10]. By employing multiple transmit antennas, Tonello [11], [12] suggested space-time BICM schemes for narrowband radio channels in flat fading cases. In addition, similar ideas have been proposed in other papers such as [13] and [14].

For wideband systems, both single carrier and multicarrier systems can be applied to combat frequency selective channels. Equalization is one choice for the single carrier approach [15]. Orthogonal frequency division multiplexing (OFDM) has been proposed for a wide range of radio channels [16]–[18]. Especially, the wireless local area network (LAN) system defined by the IEEE 802.11a standard adopts OFDM in packet-based communications operating in the unlicensed 5-GHz band. In the case of frequency-selective fading channels, Tonello proposed to use an equalizer rather than an OFDM modem [19]. In addition, in [20], space-time codes developed in [2] were combined with OFDM modulation.

In this paper, we present coded modulation schemes for space-time coding with bit-interleaved coded modulation combined with a multicarrier OFDM modulation system. By applying the well-known Turbo principle [21], [22], the decoder performance is significantly improved. As illustrated in the serially concatenated coded system configuration [23], [24], the demapper block and the convolutional decoder act as an inner and an outer code, respectively. The near optimum decoder uses a maximum *a posteriori* (MAP) decoder [25] in an iterative configuration.

One of the biggest advantage of the proposed space-time bit-interleaved coded modulation (ST-BICM) system is its flexibility in terms of various system configurations. Unlike other space-time coded schemes in [2] where coding and modulation designs were handcrafted for each system setup, a single coder in the ST-BICM system can support many different data rates. For example, in the wireless LAN system design where eight different data rate modes are defined in the 802.11a standard, the proposed ST-BICM can support all of eight modes with a single coder. This is not possible using the type of space-time coded schemes presented in [2].

In this work, we are particularly interested in slowly changing channels such as those typically experienced in the indoor environment. These are modeled by block fading channel models [26]. As subchannels in the OFDM system are normally correlated with each other, the pairwise error probability will be analyzed based on the correlated fading assumption. Extensive simulation results will be shown in the simulation section comparing the proposed ST-BICM technique with iterative decoding and a conventional single antenna system based on the 802.11a standard. The simulation results show that the iterative decoding scheme is crucial when the number of transmit antennas is greater than that of receive antennas.

The paper is organized as follows. In Section II, the system model is briefly presented. In addition, receiver and transmitter structures are discussed. In Section III, we present pairwise error probability evaluation for the system. In Section IV, simulation results for various system configurations are presented. Then, the paper is concluded in Section V.

II. System Model and Receiver/Transmitter Structures

We consider systems with $N_t \geq 2$ transmit antennas and $N_r \geq 1$ receive antennas in this paper. N_t is typically larger than or equal to N_r. As indicated above and as will be explained in detail in the next section, we will use space-time coding with bit-interleaved coded modulation. The modem constellations used are M-PSK or M-QAM with M constellation points. Furthermore, in this paper, we consider wideband channels by using OFDM rather than single carrier modems. The spectral efficiency of the system is [11] $R_T = R_C \cdot N_t \cdot \log_2 M$ bits/s/Hz, where R_C is the rate of the convolutional code used.

First, we consider an OFDM system with F subcarriers. Assuming proper cyclic prefix operation, the output at the kth subcarrier and at

Manuscript received February 27, 2003; revised April 22, 2003. This work was supported in part by a Korea University grant, in part by the Korea Science and Engineering Foundation under Grant R08-2003-000-10761-0, and in part by a University IT Research Center Project. This paper was presented in part at International Conference on Communications, May 2003. The associate editor coordinating the review of this paper and approving it for publication was Dr. Rick S. Blum.

I. Lee is with the Department of Communications Engineering, Korea University, Seoul, Korea (inkyu@korea.ac.kr).

A. M. Chan is with the Massachusetts Institute of Technology, Cambridge, MA 02139 USA (chanal@mit.edu).

C.-E. W. Sundberg is with the SundComm, Chatham, NJ 07928 USA (cews@ieee.org).

Digital Object Identifier 10.1109/TSP.2003.822350

the lth time slot from the jth receive antenna matched filter after the discrete Fourier transform (DFT) is given by

$$y_{k,l}^{j} = \sqrt{E_s}\sum_{i=1}^{N_t} H_{k,l}^{i,j} x_{k,l}^{i} + n_{k,l}^{j} \quad \text{for } j = 1, 2, \cdots N_r \tag{1}$$

where $x_{k,l}^{i}$ is the transmitted symbol at the ith transmit antenna at the kth subcarrier and at the lth time slot. E_s is the symbol energy, and $H_{k,l}^{i,j}$ is the equivalent channel frequency response of the link between the ith transmit antenna and jth receive antenna at the kth subcarrier and at the lth time slot. Finally, in (1), $n_{k,l}^{j}$ is a sequence of i.i.d. complex zero mean Gaussian variables with variance $N_0/2$ per dimension.

In this channel system, we make the following model assumptions. Considering the time domain channel impulse response between the ith transmit and jth receive antenna, a frequency-selective channel can be modeled as

$$h^{i,j}(t,\tau) = \sum_{n=1}^{K} \bar{h}^{i,j}(n;t)\delta(\tau - \tau_n)$$

where the channel coefficients $\bar{h}^{i,j}(n;t)$ are independent complex Gaussian with zero mean (Rayleigh fading), $\delta(\cdot)$ is the Dirac delta function, and K denotes the number of channel taps.

Furthermore, the channel impulse responses of the antenna links are independent of the different links. Both fast fading (i.e., uncorrelated fading coefficients in time) and block fading (i.e., static fading coefficients over a block of transmitted symbols, independent over blocks) are possibilities for this channel model. In this work, we will concentrate on the block fading model describing wireless LANs with slow movements. Therefore, we will omit the time indices l and t for simplicity.

Neglecting leakage, it follows that the channel frequency response in (1) can be expressed by

$$H_k^{i,j} = \sum_{n=1}^{K} \bar{h}^{i,j}(n) e^{-j\frac{2\pi k \tau_n}{FT}} = \bar{\mathbf{h}}_{i,j}^{*}\mathbf{w}_k \tag{2}$$

where T represents the sampling period. In addition, we denote $\bar{\mathbf{h}}_{i,j} = [\bar{h}^{i,j}(1)\bar{h}^{i,j}(2)\cdots\bar{h}^{i,j}(K)]^*$ and $\mathbf{w}_k = [e^{-j2\pi k\tau_1/FT} e^{-j2\pi k\tau_2/FT} \cdots e^{-j2\pi k\tau_K/FT}]^T$. Note that $|H_k^{i,j}|$ is Rayleigh distributed and is correlated in frequency.

The symbol constellation is normalized such that $\mathrm{E}\{|x_k^i|^2\} = 1$ for $i = 1, 2, \cdots N_t$. With vector notations, (1) can be written as

$$\mathbf{y}_k = \begin{bmatrix} y_k^1 \\ \vdots \\ y_k^{N_r} \end{bmatrix} = \sqrt{E_s} \begin{bmatrix} H_k^{1,1} & \cdots & H_k^{N_t,1} \\ \vdots & \ddots & \vdots \\ H_k^{1,N_r} & \cdots & H_k^{N_t,N_r} \end{bmatrix} \begin{bmatrix} x_k^1 \\ \vdots \\ x_k^{N_t} \end{bmatrix} + \begin{bmatrix} n_k^1 \\ \vdots \\ n_k^{N_r} \end{bmatrix}$$

or equivalently

$$\mathbf{y}_k = \sqrt{E_s}\mathbf{H}_k\mathbf{x}_k + \mathbf{n}_k, \quad \text{for } k = 1, 2, \cdots F.$$

Fig. 1 shows a transmitter structure for space-time coding with bit-interleaved coded modulation with OFDM modulation for the case of wideband frequency-selective channels. The convolutional coder is using a binary convolutional code with rate R_C. A key component in these systems is the mapper that converts b bits to modem symbols in the M-ary symbols in the sequence x_k^i mentioned above. In this structure, there are F parallel narrowband flat-fading subchannels being transmitted over each antenna.

Fig. 2 shows an iterative decoder for the space-time code for the OFDM system. For the wideband system, the demodulator is an

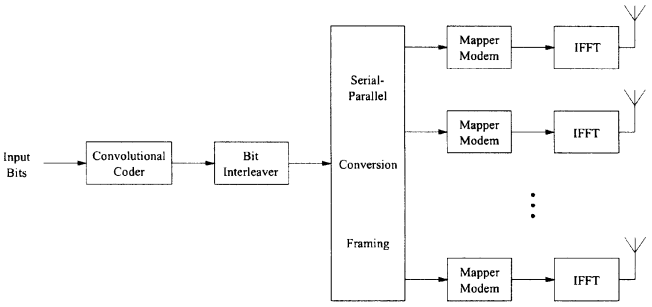

Fig. 1. Transmitter structure for wideband OFDM systems.

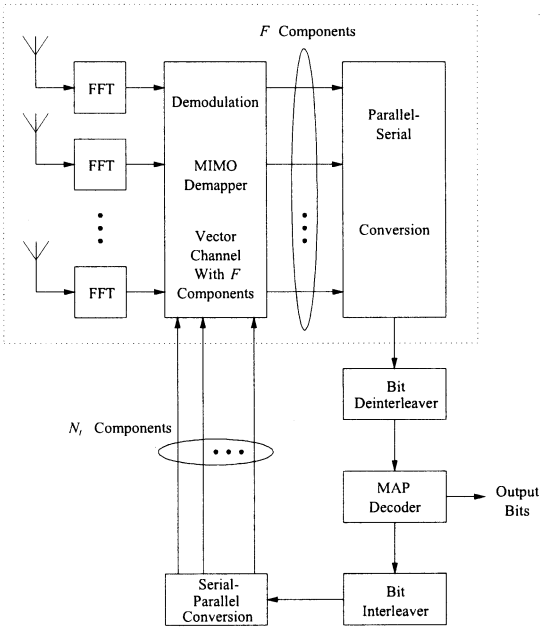

Fig. 2. Receiver structure for wideband OFDM systems.

OFDM demodulator for each antenna, and the processing is done over F parallel subchannels. Note that the receiver now has to estimate the F multi-input multi-output (MIMO) subchannels (slowly varying flat fading) rather than only one in the narrowband flat fading case. A key subsystem in the receiver is that the MIMO demapper, which is based on the N_r antenna inputs, produces the MAP estimates of the demapped bits for each subcarrier in the N_t transmitter streams corresponding to the N_t antennas in the transmitter.

Let $d_k^{i,m}$ be the bit that is mapped at the kth subcarrier into the mth bit position ($m = 1, 2, \cdots \log_2 M$) of the constellation symbol of the ith transmit antenna ($i = 1, 2, \cdots N_t$). Then, the log likelihood ratios are given by

$$L\left(d_k^{i,m}\right) = \log\frac{P\left(d_k^{i,m} = +1\right)}{P\left(d_k^{i,m} = -1\right)}. \tag{3}$$

Let the set $\mathcal{S}_d^{i,m}$, $d = +1$, or -1 be a set of all symbol vectors with a $+1$ or -1 value of bit $d_k^{i,m}$, respectively. The number of elements in such a set is $2^{N_t \log_2 M - 1}$. The log likelihood ratio (LLR) in (3) conditioned on the channel state information is

$$\log\frac{P\left(d_k^{i,m} = +1|\mathbf{y}_k, \mathbf{H}_k\right)}{P\left(d_k^{i,m} = -1|\mathbf{y}_k, \mathbf{H}_k\right)} = \log\frac{\sum_{\mathbf{x}_k \in \mathcal{S}_{+1}^{i,m}} p(\mathbf{x}_k, \mathbf{y}_k, \mathbf{H}_k)}{\sum_{\mathbf{x}_k \in \mathcal{S}_{-1}^{i,m}} p(\mathbf{x}_k, \mathbf{y}_k, \mathbf{H}_k).} \tag{4}$$

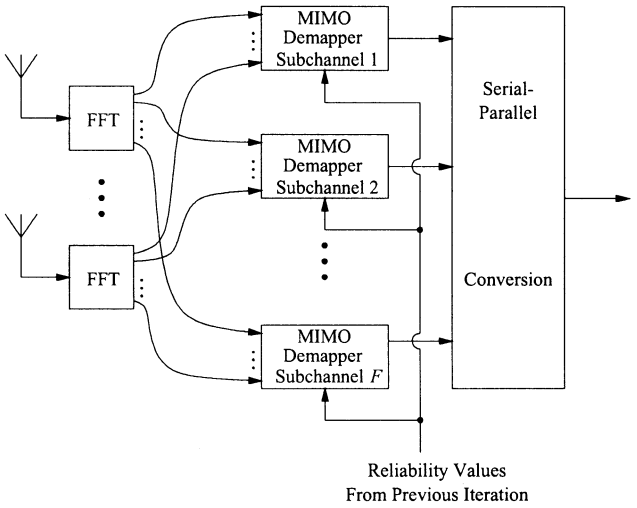

Fig. 3. MIMO demapper structure for wideband OFDM systems.

The joint probability density in (4) is then related to

$$p(\mathbf{x}_k, \mathbf{y}_k, \mathbf{H}_k) \propto \exp\left(-\frac{1}{N_0}\|\mathbf{y}_k - \sqrt{E_s}\mathbf{H}_k\mathbf{x}_k\|^2 + \frac{1}{2}\mathbf{d}_k^T\mathbf{L}_k\right)$$

where $\mathbf{d}_k$ and $\mathbf{L}_k$ are column vectors of length $N_t \cdot \log_2 M$ comprised of $d_k^{i,m}$ and $L(d_k^{i,m})$, respectively.

The elements of the *a priori* likelihood ratio vector $\mathbf{L}_k$ are assumed to be independent in the interleaved bit stream. For the first pass through the iterative decoder, these are set to zero. For the second pass and beyond, they are derived from the MAP decoder for the convolutional decoder in Fig. 2. Then, by subtracting the input likelihood values from the output LLR values computed in (4), the extrinsic information is exchanged between the MIMO demapper and the MAP decoder to improve the bit error rate (BER) performance.

The MIMO demapper shown inside a dotted line box in Fig. 2 is described in more detail in Fig. 3. The signal outputs of the kth subcarrier for all N_r receive antennas are sent to a MIMO demapper box for subchannel k, which computes the log likelihood ratio using (4), combining a likelihood ratio vector $\mathbf{L}_k$ from the previous iteration.

III. Pairwise Error Probability Evaluation

In this section, we will evaluate the pairwise error probability (PEP) for the correlated fading case. Consider the block fading frequency-selective channel with F carrier OFDM system assuming constant fading over one OFDM symbol. It has been shown in [10] and [11] that the performance of the iterative decoding scheme approaches the maximum likelihood (ML) performance with the assumption of exact feedback from the decoder. Assuming that $\mathbf{x}$ is transmitted, we have the average PEP that the ML chooses $\hat{\mathbf{x}}$ over $\mathbf{x}$ as

$$P\left(\mathbf{x} \to \hat{\mathbf{x}}|\{\mathbf{H}_k\}_{k=1}^F\right) = P\left(\sum_{k=1}^F \left(\|\mathbf{y}_k - \sqrt{E_s}\mathbf{H}_k\mathbf{x}_k\|^2 - \|\mathbf{y}_k - \sqrt{E_s}\mathbf{H}_k\hat{\mathbf{x}}_k\|^2\right) > 0|\{\mathbf{H}_k\}_{k=1}^F\right).$$

Using a Chernoff bound, this PEP is bounded by

$$P\left(\mathbf{x} \to \hat{\mathbf{x}}|\{\mathbf{H}_k\}_{k=1}^F\right) \le \exp\left(-\frac{E_s}{4N_0}\sum_{k=1}^F \|\mathbf{H}_k\mathbf{e}_k\|^2\right) \quad (5)$$

where $\mathbf{e}_k$ is defined as $\mathbf{e}_k = \hat{\mathbf{x}}_k - \mathbf{x}_k$.

It is well known that Chernoff bounds are tight in the exponential sense for high signal-to-noise ratio (SNR). However, the bound is off approximately by a multiplicative constant factor, which is independent of SNR. Here, we are primarily interested in the slope of the error probability and the code parameters that affect it. This will give an indication on how to design the codes to obtain a large degree of diversity in the system. More precise error probability curves are obtained by simulations below.

In order to compute the above bound, let us first denote the jth row of $\mathbf{H}_k$ as $\mathbf{h}_{j,k}^*$. Then, using (2), $\mathbf{h}_{j,k}^*$ can be expressed as

$$\mathbf{h}_{j,k}^* = \left[\bar{\mathbf{h}}_{1,j}^*\bar{\mathbf{h}}_{2,j}^* \cdots \bar{\mathbf{h}}_{N_t,j}^*\right]\begin{bmatrix}\mathbf{w}_k & 0 & \cdots & 0\\ 0 & \mathbf{w}_k & \ddots & \vdots\\ \vdots & \ddots & \ddots & 0\\ 0 & \cdots & 0 & \mathbf{w}_k\end{bmatrix} = \tilde{\mathbf{h}}_j^*\mathbf{W}_k^*$$

where $\tilde{\mathbf{h}}_j$ is a column vector of length KN_t, and $\mathbf{W}_k$ is an N_t by KN_t matrix. Note that elements of $\tilde{\mathbf{h}}_j$ are uncorrelated with each other and are independent of k in frequency.

Defining $\tilde{\mathbf{e}}_k = \mathbf{W}_k^*\mathbf{e}_k$, we can express $\sum_{k=1}^F \|\mathbf{H}_k\mathbf{e}_k\|^2$ in (5) as

$$\sum_{k=1}^F \|\mathbf{H}_k\mathbf{e}_k\|^2 = \sum_{k=1}^F\sum_{j=1}^{N_r}\left|\mathbf{h}_{j,k}^*\mathbf{e}_k\right|^2 = \sum_{k=1}^F\sum_{j=1}^{N_r}\left|\tilde{\mathbf{h}}_j^*\tilde{\mathbf{e}}_k\right|^2 = \sum_{k=1}^F\sum_{j=1}^{N_r}\tilde{\mathbf{h}}_j^*\tilde{\mathbf{E}}_k\tilde{\mathbf{h}}_j = \sum_{j=1}^{N_r}\tilde{\mathbf{h}}_j^*\mathbf{E}\tilde{\mathbf{h}}_j \quad (6)$$

where $\tilde{\mathbf{E}}_k$ and $\mathbf{E}$ are denoted by $\tilde{\mathbf{E}}_k = \tilde{\mathbf{e}}_k\tilde{\mathbf{e}}_k^*$ and $\mathbf{E} = \sum_{k=1}^F \tilde{\mathbf{E}}_k$, respectively.

Since $\mathbf{E}$ is a non-negative definite Hermitian matrix, we have an eigendecomposition as $\mathbf{E} = \mathbf{P}^*\mathbf{\Lambda}\mathbf{P}$, where $\mathbf{P}$ is a unitary matrix, and $\mathbf{\Lambda}$ is a real diagonal matrix. Substituting this into (6), we now get

$$\sum_{k=1}^F \|\mathbf{H}_k\mathbf{e}_k\|^2 = \sum_{j=1}^{N_r}\hat{\mathbf{h}}_j^*\mathbf{\Lambda}\hat{\mathbf{h}}_j$$

where $\hat{\mathbf{h}}_j = \mathbf{P}\tilde{\mathbf{h}}_j$ is a column vector of length KN_t.

Defining λ_i as the ith eigenvalue of $\mathbf{E}$ finally yields

$$\sum_{k=1}^F \|\mathbf{H}_k\mathbf{e}_k\|^2 = \sum_{i=1}^{KN_t}\sum_{j=1}^{N_r}\lambda_i\left|\hat{H}_j^i\right|^2 \quad (7)$$

where $\hat{H}_j^i$ is the ith element of $\hat{\mathbf{h}}_j$.

Since $\mathbf{P}$ is unitary, $|\hat{H}_j^i|$ is also independent and Rayleigh distributed. Therefore, substituting (7) into the bound (5) and averaging it with respect to independent Rayleigh distributions of $|\hat{H}_j^i|$ yields

$$P(\mathbf{x} \to \hat{\mathbf{x}}) \le \prod_{i=1}^{KN_t}\prod_{j=1}^{N_r}\frac{1}{1 + \frac{E_s}{4N_0}\lambda_i}.$$

Here, we assume that $\hat{H}_j^i$ has equal power for all i and j to simplify the analysis.

Let R denote the rank of matrix $\mathbf{E}$. Then, $KN_t - R$ eigenvalues of $\mathbf{E}$ are zero. Let us define the nonzero eigenvalues of $\mathbf{E}$ as $\lambda_1, \lambda_2, \cdots \lambda_R$. Finally, it follows that at high SNR, the PEP bound is given by

$$P(\mathbf{x} \to \hat{\mathbf{x}}) \le \prod_{i=1}^{R}\left(\frac{E_s}{4N_0}\lambda_i\right)^{-N_r}. \quad (8)$$

We now see that the diversity order given by the above bound is determined by RN_r. Next, we further examine conditions on the rank R. Let us denote L as the number of instances $1 \le k \le F$ such that $\mathbf{e}_k\mathbf{e}_k^*$ is nonzero. Note that L is related to the effective length of the code, as defined in [7]. Considering that $\tilde{\mathbf{e}}_k\tilde{\mathbf{e}}_k^*$ is a rank one matrix, we can show that the rank R of $\mathbf{E} = \sum_{k=1}^F \tilde{\mathbf{e}}_k\tilde{\mathbf{e}}_k^*$ is obtained by $\min(L, KN_t)$. Therefore, it can be seen here that the maximum achievable diversity order of the ST-BICM system for K tap equal power fading channels is KN_tN_r. Note that this number is K times higher than the full diversity order N_tN_r in flat fading cases [2].

Here, we consider two cases: $L > KN_t$ and $L \leq KN_t$. The first case is encountered when the channel does not contain enough diversity in frequency. In this case, as long as L is greater than KN_t, which is determined by the channel, the system performance is mostly limited by the system parameter such as N_t. Therefore, we expect a large improvement in performance by the multiple transmit antenna for channels with small K.

The second case accounts for the situation when the channel delay spread is large. As an extreme case, the channel frequency responses $H_k^{i,j}$ at different subcarriers become fully uncorrelated. This occurs when $K = F$ and all the amplitude taps $\bar{h}^{i,j}(n)$ in (2) have equal variance. In this case, the diversity order of the system is determined by LN_r, which is independent of the number of transmit antenna N_t. Rather, the number of nonzero elements $\mathbf{e}_k\mathbf{e}_k^*$ specified by the code design determines the diversity order. A similar derivation was made in the space-time coded system with low-density parity check (LDPC) code, which was proposed in [27].

As we assume that the channel information is available only to the receiver side, it is not feasible to design a code that can maximize the system performance for every channel delay profiles. In general, the derived bound leads to a code design consideration for the ST-BICM system in OFDM: First, for the channels with small K, it is important to have a code whose effective length L is greater than KN_t in order to achieve the full diversity order KN_t. Second, for the channels with large K, which is observed in many indoor wireless LAN systems, the diversity order is maximized by employing a standard maximum Hamming distance convolutional code. It should be noted that in comparison with the single antenna case, the slope change in BER curves due to additional transmit antennas may not be as distinct as that observed in channels with a small delay spread. Therefore, optimizing the coding gain becomes more important than maximizing the diversity order, as the diversity order exploited by the single antenna system may already be big enough when channels contain large delay spreads. Then, one should choose an appropriate mapping and coding to maximize the coding gain which is determined by $(\prod_{i=1}^{R}\lambda_i)^{-N_r}$. One way of maximizing this coding gain is to increase the effective length of the code, as suggested in [2]. In addition, employing an interleaver is another important way of improving the coding gain, as the interleaver can spread out the error sequences $\mathbf{e}_k\mathbf{e}_k^*$. Note that as $\mathbf{E}$ changes with the channel delay profile, optimizing $\prod_{i=1}^{R}\lambda_i$ for all possible channel realizations is not possible.

From the theory above, we can now construct space-time coding schemes at different levels of diversity degree and different levels of bandwidth efficiency. Given the overall decoder complexity, we can optimize a convolutional code with the maximum Hamming distance by simply increasing the code memory. However, it should be noted that decoder complexity grows exponentially with code memory.

Note that unlike other space-time coding schemes based on trellis codes such as the space-time trellis codes (STTC) proposed in [2], the trellis complexity of the ST-BICM for the given spectral efficiency is independent of the modulation level and the number of transmit antennas. In addition, by applying a puncturing mechanism in convolutional codes [28], it is possible to achieve a high spectral efficiency without employing a high modulation level, which makes a receiver design difficult. Thus, by using smaller signal constellations and higher rate codes at a given N_t, lower complexity receivers are achieved. As a result, the ST-BICM schemes can obtain a spectral efficiency higher than 4 bits/s/Hz without difficulties, whereas it is impractical for STTC in [2] to achieve such a high spectral efficiency. This is quite important, as the target spectral efficiency in future wireless systems becomes higher. Flexibility of the ST-BICM scheme is also critical in many practical wireless systems where different data rates need to be supported. For example, for the 802.11a standard, there should be eight different decoders with the STTC in [2], whereas the ST-BICM utilizes a single binary convolutional decoder with different puncturing patterns.

TABLE I
MODULATION LEVEL AND CODE POLYNOMIALS

	R_T=2 bps/Hz	R_T=4.5 bps/Hz
802.11a	16QAM	64QAM
	$R_C = 1/2$ (133,171) code	$R_C = 3/4$ (133,171) code
ST-BICM	QPSK	8PSK
($N_t = 2$)	$R_C = 1/2$ (23,35) code	$R_C = 3/4$ (23,35) code

IV. SYSTEM SIMULATIONS

In this section, we present simulation results for the MIMO ST-BICM in the indoor wireless channel environment and compare them with a single antenna system defined in the 802.11a wireless LAN standard. In the simulations, we compare both systems in 24- and 54-Mbits/s modes, which is the maximum throughput supported by the 802.11a system. The system parameters are shown in Table I. The binary convolutional code polynomials are represented in octal notation. The spectral efficiencies for each mode, excluding overhead due to guard bands and guard symbols, are 2 and 4.5 bits/s/Hz, respectively.

First, we present the ST-BICM system performance with the iterative decoding scheme. To make a fair comparison, both the MIMO ST-BICM and the 802.11a systems are set to have the same spectral efficiency. For practical implementation issues, two transmit antennas with one and two receive antenna systems are considered. The indoor wireless channel is considered quasistatic as mobility of users is quite low. We use a typical indoor channel model with 5-tap power delay profile ($K = 5$) having an exponentially decayed fading characteristics, and each ray is assumed to be independently Rayleigh fading.

The OFDM modulation defined in the 802.11a standard specifies 64 point FFT. One OFDM symbol duration is 4 μs, including the 0.8-μs guard interval. This specification is designed to handle the root means square (RMS) delay spread of about 250 ns. The 5-tap multipath channel with exponentially decayed delay profile accounts for approximately the RMS spread of 100 ns. Throughout the simulations, one packet is assumed to consist of one OFDM symbol for simplicity. A random interleaver is used for simulations, and its size is determined by $F \cdot N_t \cdot \log_2 M$. This short length was chosen because of latency requirement. Longer interleavers are expected to yield somewhat improved performance. Throughout the paper, Gray mapping is employed in all system simulation cases. The number of the decoding iteration is set to 5. However, it is noted from the simulations that after three iterations the BER performance is normally saturated. As for the code choice, we use the 16 state memory 4 (23,35) code, if not specified otherwise.

First, to demonstrate the validity of the PEP analysis presented in the previous section, the simulation results with a flat fading channel are presented in Fig. 4. The x-axis represents the received SNR per antenna in decibels, and the y-axis indicates the frame error rate (FER). In this case, the frame length is set to 128 bits. It is clear from this plot that 2 by 1 and 2 by 2 systems exhibit the diversity gain slope of 2 and 4, respectively, as shown in [11]. This also matches well with the PEP analysis derived in the previous section for the $K = 1$ case. Compared with these flat fading cases, the frequency-selective channel already contains some level of frequency diversity inherent in the channel. Therefore, we do not see as distinctive slope changes for the frequency-selective channel as for the flat fading channel.

Figs. 5 and 6 show simulation results for 2- and 4.5-bits/s/Hz cases, respectively, for the 5-tap exponentially faded channel model. When compared with the flat fading case in Fig. 4, the overall FER curve

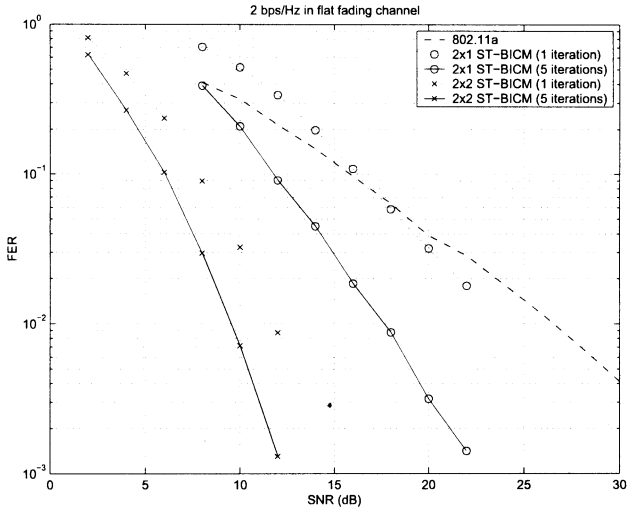

Fig. 4. Comparison in flat fading channels.

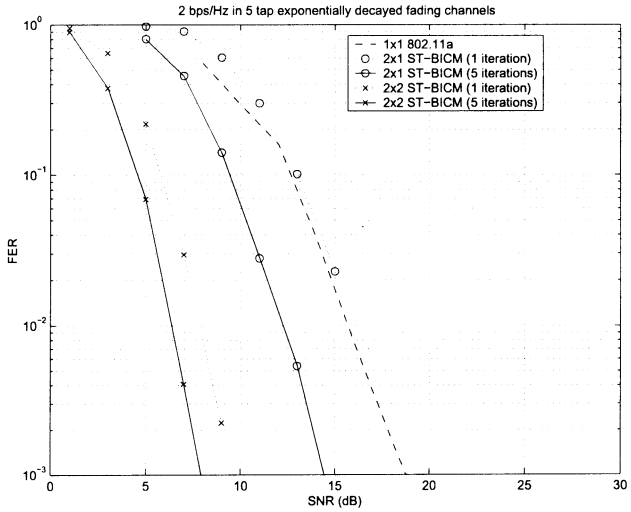

Fig. 5. Comparison for 2 bits/s/Hz in 5-tap exponentially decayed fading channel.

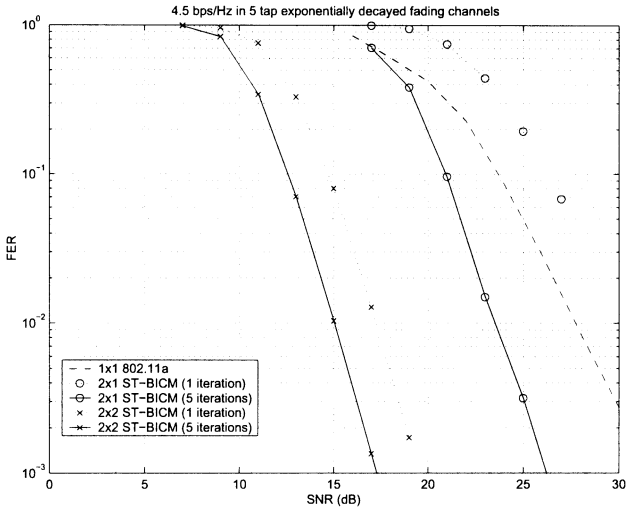

Fig. 6. Comparison for 4.5 bits/s/Hz in 5-tap exponentially decayed fading channel.

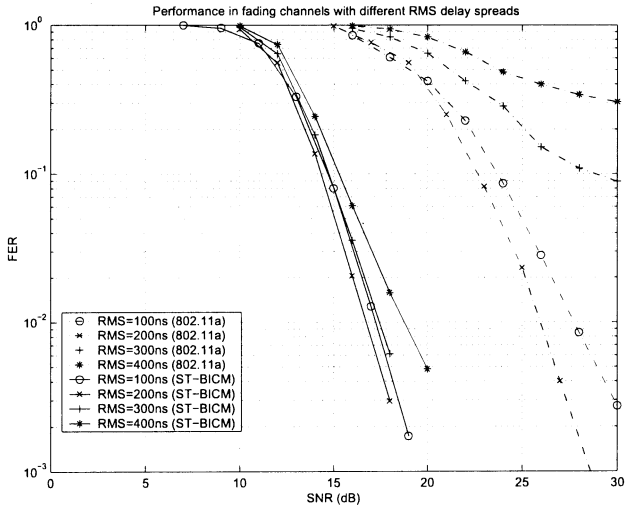

Fig. 7. Comparison of 802.11a system and the 2 by 2 ST-BICM without iterations for 4.5 bits/s/Hz in exponentially decayed fading channels with various RMS delay spreads.

slopes are steeper due to the frequency-selective nature of the channel. Actually, it appears that there are no distinctive differences in the FER curve slopes between one and two transmit antennas. This confirms the observation made in the analysis section that the slope change due to additional transmit antennas may become indistinguishable when the channel already contains enough diversity in frequency. As a result, the performance gain of the 2 by 1 system over the single antenna case reduces to 5 dB at 1% FER. However, the 2 by 2 system still exhibits a gain of 9 dB over the single antenna case at 2 bits/s/Hz, and the gain increases to 12.5 dB for the 4.5-bits/s/Hz case.

One interesting thing we observe in these plots is that in the 2 by 1 system, the performance with five iterations is much better than that with one iteration, whereas in the 2 by 2 system, the performance gain is only 2 dB with five iterations. The reason for this is that the MIMO demapper in the 2 by 1 system needs to deal with one equation y_k^1 in (4), which contains two unknown variables x_k^1 and x_k^2. Therefore, the performance at the MIMO demapper with no *a priori* information at the first iteration is very poor, and this is greatly improved by the subsequent iterations. In contrast, the 2 by 2 system can generate relatively reliable output values at the MIMO demapper. Thus, the performance with one iteration is already close to that with many iterations.

Rather poor performance with one iteration when the system employs more transmit antennas than receive antennas is then attributed to inaccurate demodulation operations in the demapper block, which cannot be improved without an aid of iterative decoding. Similarly, it has been observed in [29], which did not consider an iterative decoder, that the use of $N_t > N_r$ leads to poor performance in fast fading cases. This indicates that it is crucial to apply iterative decoding schemes for the ST-BICM system when $N_t > N_r$.

In the 2 by 2 system, we may resort to one iteration as observed in the previous section. This observation leads to a much simpler architecture for the 2 by 2 system. If we take only one iteration, which means no feedback from the decoder, then the MAP decoder can be replaced by a simple Viterbi decoder. Therefore, the overall complexity would be considerably reduced at the expense of only 2-dB performance loss compared with the case with five iterations. In addition, this structure allows a low latency, which is important to some delay-sensitive applications such as real time transmission.

Therefore, from now on, we focus on the 2 by 2 ST-BICM system with a conventional Viterbi decoder replacing a MAP decoder without any iterations. Fig. 7 presents the performance comparison between the 2 by 2 system with the 802.11a standard for 4.5-bits/s/Hz system

in exponentially decayed fading channels with various RMS delay spreads. Note that for exponentially faded channels, the exponential decay constant is equal to the RMS delay spread. This plot shows that the 802.11a system distinctively exhibits the irreducible error floor in channels with the RMS delay spread exceeding 300 ns. Compared with this, the 2 by 2 ST-BICM system shows quite robust system performance, even without iterative decoding schemes in channels with RMS delay spreads as high as 400 ns. It is clear from the plot that compared with the conventional single antenna system, the proposed noniterative 2 by 2 ST-BICM system can operate in channels with much higher delay spreads. Therefore, the guard time interval in one OFDM symbol can be reduced significantly in the 2 by 2 ST-BICM system, and this leads to a higher throughput.

It should be noted that when operated in a noniterative fashion, the baseband operation of the 2 by 2 ST-BICM system can be implemented with a lower complexity than that of the 802.11a system since the number of states in the Viterbi decoder trellis is reduced from 64 to 16. To achieve the same spectral efficiency of 4.5 bits/s/Hz, the modulation scheme of the 2 by 2 system becomes 8 PSK, which is much simpler for front-end receiver design than 64 QAM in the 802.11a standard. In addition, note that the receiver does not need to estimate the noise power N_0 because by max-log approximation, the LLR computation in (4) is simplified to

$$\min_{\mathbf{x}_k \in \mathcal{S}_{+1}^{i,m}} \|\mathbf{y}_k - \sqrt{E_s}\mathbf{H}_k\mathbf{x}_k\|^2 - \min_{\mathbf{x}_k \in \mathcal{S}_{-1}^{i,m}} \|\mathbf{y}_k - \sqrt{E_s}\mathbf{H}_k\mathbf{x}_k\|^2.$$

The only complexity increase in the baseband operation comes from the above demapper block computation of $\|\mathbf{y}_k - \sqrt{E_s}\mathbf{H}_k\mathbf{x}_k\|^2$, as it now requires 20 multiplications instead of 6. However, the number of candidates to compute the optimum demapper values remains the same for both systems in one demapper block. Therefore, the total complexity in terms of baseband operation in the proposed 2 by 2 system can be lower than the conventional system.

V. Discussion and Conclusions

A flexible class of space-time codes based on bit-interleaved coded modulation systems has been derived and evaluated for wideband frequency-selective channels. We have also analyzed the pairwise error probability for the OFDM system with multiple antennas assuming correlated fading channels. Computer simulations demonstrate the efficacy of the systems for block fading channels using an OFDM system. One possible application is future wireless LANs. We have found that an iterative decoding scheme is crucial when there are more transmit antennas than receive antennas. It has also been shown that a 2 by 2 system without iterations can achieve more than 10 dB gain over a single antenna system, and it exhibits robust performance for channels with severe delay spreads. Moreover, the baseband operation complexity decreases in comparison to the 802.11a standard. Many issues remain to be looked into, especially the choice of mapper and interleaver for the best engineered systems.

References

[1] J. G. Proakis, *Digital Communications*. New York: McGraw-Hill, 1983.

[2] 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, pp. 744–765, Mar. 1998.

[3] S. L. Ariyavisitakul, "Turbo space-time processing to improve wireless channel capacity," *IEEE Trans. Commun.*, vol. 48, pp. 1347–1359, Aug. 2000.

[4] G. J. Foschini, "Layered space-time architecture for wireless communication in a fading environment when using multi-element antennas," *Bell Labs. Tech. J.*, vol. 1, pp. 41–59, 1996.

[5] I. E. Telatar, "Capacity of multi-antenna Gaussian channels," *Eur. Trans. Telecom.*, vol. 10, pp. 585–595, Nov. 1999.

[6] S. M. Alamouti, "A simple transmitter diversity scheme for wireless communications," *IEEE J. Select. Areas Commun.*, pp. 1451–1458, Oct. 1998.

[7] C. Schlegel and D. J. Costello, "Bandwidth efficient coding for fading channels: Code construction and performance analysis," *IEEE J. Select. Areas Commun.*, vol. 7, pp. 1356–1368, Dec. 1989.

[8] D. Divsalar and M. K. Simon, "The design of trellis coded MPSK for fading channels: performance criteria," *IEEE Trans. Commun.*, vol. 36, pp. 1004–1012, Sept. 1988.

[9] E. Zehavi, "8-PSK trellis codes for a Rayleigh channel," *IEEE Trans. Commun.*, vol. 40, pp. 873–884, May 1992.

[10] X. Li and J. A. Ritcey, "Trellis-coded modulation with bit interleaving and iterative decoding," *IEEE J. Select. Areas Commun.*, vol. 17, pp. 715–724, Apr. 1999.

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

[12] ——, "Performance of space-time bit-interleaved codes in fading channels with simplified iterative decoding," in *Proc. of IEEE Veh. Technol. Conf.*, Sept. 2001, pp. 1357–1361.

[13] E. Biglieri, G. Taricco, and E. Viterbo, "Bit-interleaved time-space codes for fading channels," in *Proc. Conf. Inform. Sci. Syst.*, 2000, pp. WA4.1–WA4.6.

[14] B. Hochwald and S. T. Brink, "Achieving near-capacity on a multiple-antenna channel," *IEEE Trans. Commun.*, vol. 51, pp. 389–399, Mar. 2003.

[15] S. L. Ariyavisitakul, J. H. Winters, and I. Lee, "Optimum space time processors with dispersive interference: unified analysis and required filter span," *IEEE Trans. Commun.*, vol. 57, pp. 1073–1083, July 1999.

[16] L. J. Cimini, "Analysis and simulation of a digital mobil channel using orthogonal frquency division multiplexing," *IEEE Trans. Commun.*, vol. COM-33, pp. 665–675, July 1985.

[17] H. Sari, G. Karam, and I. Jeanclaude, "Transmission techniques for digital terrestrial TV broadcasting," *IEEE Commun. Mag.*, pp. 100–109, Feb. 1995.

[18] R. van Nee, G. Awater, M. Morikura, H. Takanashi, M. Webster, and K. W. Halford, "New high-rate wireless LAN standards," *IEEE Commun. Mag.*, vol. 37, pp. 82–88, Dec. 1999.

[19] A. M. Tonello, "Space-time bit-interleaved coded modulation over frequency selective fading channels with iterative decoding," in *Proc. IEEE GLOBECOM*, Nov. 2000, pp. 1616–1620.

[20] B. Lu and X. Wang, "Space-time code design in OFDM systems," in *Proc. IEEE GLOBECOM*, Nov. 2000, pp. 1000–1004.

[21] C. Berrou, A. Glavieux, and P. Thitimajshima, "Near Shannon limit error-correcting coding and decoding: Turbo-codes," in *Proc. IEEE Int. Conf. Commun.*, May 1993, pp. 1064–1070.

[22] J. Hagenauer, "The turbo principle: Tutorial introduction and state of the art," in *Proc. Int. Symp. Turbo Codes*, 1997, pp. 1–11.

[23] S. Benedetto, D. Divsalar, G. Montorsi, and F. Pollara, "Serial concatenation of interleaved codes: Performance analysis, design, and iterative decoding," *IEEE Trans. Inform. Theory*, vol. 45, pp. 909–926, May 1998.

[24] I. Lee, "The effect of a precoder on serially concatenated coding systems with ISI channel," *IEEE Trans. Commun.*, vol. 49, pp. 1168–1175, July 2001.

[25] L. Bahl, J. Cocke, F. Jelinek, and J. Raviv, "Optimal decoding of linear codes for minimizing symbol error rate," *IEEE Trans. Inform. Theory*, vol. IT-20, pp. 284–287, Mar. 1974.

[26] R. Knopp and P. A. Humblet, "On coding for block fading channels," *IEEE Trans. Inform. Theory*, vol. 46, pp. 189–205, Jan. 2000.

[27] 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, pp. 74–88, Jan. 2002.

[28] D. Haccoun and G. Begin, "High-rate punctured convolutional codes for Viterbi and sequential decoding," *IEEE Trans. Commun.*, vol. 38, pp. 1113–1125, Nov. 1989.

[29] S. H. Müller-Weinfurtner, "Coding approaches for multiple antenna transmission in fast fading and OFDM," *IEEE Trans. Signal Processing*, vol. 50, pp. 2442–2450, Oct. 2002.