On the Design of Layered Space-Time Systems for Autocoding

Hesham El Gamal, *Member, IEEE*

***Abstract*—Recently, Hochwald *et al.* have recognized that arbitrarily reliable communication is possible in multiantenna systems with coding over only a single coherence interval. In particular, they showed that reliable communication is possible for all rates $R \leq C_a$ with code words that extend over a single coherence interval when the number of transmit antennas and coherence interval $(n, T) \rightarrow \infty$. They coined the names "*autocoding*" for this phenomenon and "*autocapacity*" for C_a. They also proposed a signalling scheme based on random unitary matrices that achieves a significant fraction of this capacity. The main limitation, however, is that currently no decoder of reasonable complexity is known for this signalling scheme. In this paper, we investigate the application of space-time layering to autocoding. We show that properly constructed layered systems can achieve the autocapacity with a reasonable complexity receiver composed of minimum mean-square error (MMSE) decision feedback multiuser detectors and single user decoders. In addition to this asymptotic result, we propose a specific layering approach, the threaded space-time layering, that combines generalized bit interleaved space-time coded modulation, iterative signal processing and pilot symbol assisted channel estimation. We show that this approach is well suited for practical systems with limited numbers of transmit antennas and small coherence intervals. Finally, we report simulation results that demonstrate the ability of the threaded approach to achieve significant fractions of the autocapacity with a realizable receiver. The simulation results also indicate significant performance gains over the recently proposed Cayley differential space-time signalling scheme in certain scenarios.**

***Index Terms*—Algebraic space-time codes, autocoding, bit interleaved coded modulation, diversity, fading channels, multiple transmit and receive antennas, space-time layering.**

I. Introduction

RECENT WORKS have investigated the information theoretic capacity of multiple antenna systems in Rayleigh fading channels where it was shown that spatial diversity allows for significantly higher transmission rates than that possible in single antenna systems. Most of these works focused on scenarios where the coding interval is smaller than the coherence time of the channel. Earlier works have used the notion of *outage capacity* to characterize the possible performance in multiple-input multiple-output (MIMO) Rayleigh fading channels [1]. The outage capacity involves a tradeoff between the transmission rate and outage probability based on the distribution of the instantaneous mutual information. If the transmission rate is higher than the instantaneous mutual information, an outage is declared because the channel cannot support this rate. In most practical fading channels, zero outage probability can be realized only by setting the transmission rate to zero.

Recently, Hochwald *et al.* proved that arbitrarily reliable communication is possible in these channels within a single coherence interval if the duration of the coherence interval and the number of transmit antennas are allowed to grow without limit [2]. More precisely, they showed that for a fixed number of receive antennas m and a fixed total signal-to-noise ratio (SNR), there exists a code that achieves a zero block probability of error as the number of transmit antennas and coherence interval $(n, T) \rightarrow \infty$ if the transmission rate $R \leq C_a$, where $C_a = m \log_2(1 + \mathrm{SNR})$. They coined the names "*autocapacity*" for C_a and "*autocoding*" for this phenomenon. Since the coding interval is restricted to be smaller than the coherence interval in autocoding, achieving the autocapcity relies only on spatial diversity and the $T \times n$ matrix-valued signals act as their own channel codes. Hochwald *et al.* also presented a signalling construction based on random unitary matrices that achieves a significant faction of the autocapcity [2], [3].

As pointed out in [2], the importance of the autocoding phenomenon is emphasized in scenarios that require very high quality of service in slow fading environments. It also applies to the more realistic case where the channel state information is **not** available *a priori* at the receiver. The utility of these results is, however, limited by the prohibitive exponential complexity required for maximum-likelihood decoding of the unitary signalling scheme in [3]. In this paper, we attempt to address this issue and propose a low complexity approach that benefits from our earlier work on generalizing the layered space-time architecture [4].

The layered space-time approach was proposed by Foschini in an attempt to realize the outage capacity [5] where he argued that this architecture can asymptotically achieve a lower bound on the outage capacity with reasonable receiver complexity. Here, we first show that there exists a space-time layering architecture that achieves the autocoding capacity while avoiding the exponentially growing receiver complexity. Then we propose a specific type of layering, threaded space-time layering [4], with algebraic space-time component codes and iterative signal processing for practical systems with limited numbers of transmit antennas and finite coherence intervals. We present simulation results for the proposed threaded approach that compare favorably with the upper bounds of the random unitary signal construction in [3], however, with a realizable receiver.

Paper approved by X. Wang, the Editor for Equalization of the IEEE Communications Society. Manuscript received July 2, 2001; revised November 10, 2001. This work was supported by the National Science Foundation under Grant 0118859.

The author is with the Electrical Engineering Department, the Ohio State University, Columbus, OH 43210 USA (email: helgamal@ee.eng.ohio-state.edu).

Publisher Item Identifier 10.1109/TCOMM.2002.802560.

We also compare our approach with the recently proposed Cayley differential space-time signalling scheme in certain scenarios [6].

The rest of this paper is organized as follows. Section II outlines the MIMO system model where most of the notation is borrowed from [4]. In Section III, we present the asymptotic result which shows that properly constructed layered space-time systems can achieve the autocapacity. Section IV considers the design of threaded space-time systems in practical scenarios with limited numbers of transmit antennas and finite coherence intervals. Simulation results that demonstrate the excellent performance of the proposed threaded layering approach are presented in Section V. Finally, we offer some concluding remarks in Section VI.

II. System Model

We consider a multiple antenna communication system with n transmit and m receive antennas. In this paper, we are interested in the scenario where the fading channel is frequency nonselective and the channel state information is not available at the transmitter. Fig. 1 illustrates the space-time transmitter where the channel encoder accepts input from the information source and outputs a coded stream of higher redundancy suitable for error correction processing at the receiver. The encoded output stream is modulated and distributed among the n antennas via the two dimensional spatial modulator. The transmissions from each of the n transmit antennas are simultaneous and synchronous. The signal received at each antenna is, therefore, a superposition of the n transmitted signals corrupted by additive white Gaussian noise and multiplicative fading. At the receiver end, the signal r_t^j received by antenna j at time t is given by

$$r_t^j = \sqrt{E_s} \sum_{i=1}^{n} \alpha_t^{(ij)} c_t^i + n_t^j \tag{1}$$

where E_s is the energy per transmitted symbol; $\alpha_t^{(ij)}$ is the complex path gain from transmit antenna i to receive antenna j at time t; c_t^i is the symbol transmitted from antenna i at time t; n_t^j is the additive white Gaussian noise sample for receive antenna j at time t. The noise samples are independent samples of circularly symmetric zero-mean complex Gaussian random variables with variance $N_0/2$ per dimension. At each time t, the different path gains $\alpha_t^{(ij)}$ are assumed to have a complex Gaussian distribution with zero mean and unit variance and to be statistically independent. Similar to [2], the path gains are assumed to be constant across one code word and change independently from one code word to the next. Hence, the model reduces to the quasistatic fading model studied extensively (e.g., [7], [8]).

The received signal can be expressed in vector notation as

$$\underline{r}_t = S\underline{c}_t + \underline{n}_t \tag{2}$$

where $\underline{r}_t$ is the $m \times 1$ received vector at time t; S is the $m \times n$ complex channel matrix whose ith column corresponds to the path gains for the ith transmit antenna; $\underline{c}_t$ is the $n \times 1$ transmitted vector at time t; $\underline{n}_t$ is the $m \times 1$ white Gaussian noise vector.

Without loss of generality, the code C in Fig. 1 is assumed to be defined over the discrete symbol alphabet $\mathcal{Y}$. Usually, the number of code words in C is a power of the alphabet size $|C| = |\mathcal{Y}|^k$, so that there is a one-to-one mapping $\gamma : \mathcal{Y}^k \to C$ of information k-tuples onto code words of length N. The mapping γ is an encoder for C. In this paper, we will be primarily interested in the case in which C is a binary linear code (*i.e.,* $\mathcal{Y}$ is the elementary binary field $\mathbb{F} = GF(2)$ and C is linear).

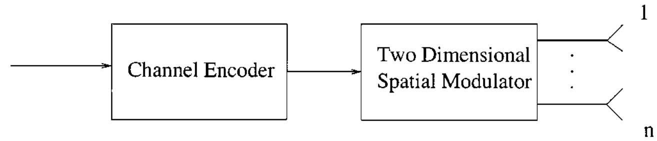


Fig. 1. Block diagram of the space-time transmitter.

The baseband modulation mapping $\mu : \mathcal{Y}^b \to \Omega$ assigns to each b-tuple of alphabet symbols a unique point in the discrete, complex-valued signalling constellation Ω, which is assumed not to contain the point zero. By extension, $\mu(\underline{v})$ denotes the modulated version of the vector $\underline{v} \in \mathcal{Y}^N$. In this case, it is understood that N must be a multiple of b and that the blocking of symbols into b-tuples for the modulator is performed left to right.

Let $\Omega^* = \Omega \cup \{0\}$ denote the expanded constellation. Then, the spatial modulator is a mapping $\mathbf{f} : \mathcal{Y}^N \to (\Omega^*)^{n \times T}$ that sends the vector $\underline{v}$ to an $n \times T$ complex-valued matrix $\mathbf{c} = \mathbf{f}(\underline{v})$, whose nonzero entries are a rearrangement of the entries of $\mu(\underline{v})$. Specifically, $\mathbf{c}$ is the baseband version of the code word $\underline{v}$ as transmitted across the channel. Thus, in the notation of (1), the matrix $\mathbf{c}$ has (i,t)th entry equal to c_t^i. Note that, in this formulation, it is expressly allowed that a complex zero (i.e., no transmission) be assigned to a given antenna at a given signalling interval; thus, $N/b \leq nT$. We will refer to n and T, respectively, as the spatial span and temporal span of $\mathbf{f}$.

III. Space-Time Layering

In [5], Foschini proposed the layered transmission approach and argued that it can realize a significant portion of the MIMO outage capacity with reasonable receiver complexity. In the layered space-time architecture, the channel encoder of Fig. 1 is composite[1] and the multiple, independent coded streams are distributed throughout the transmission resource array in so-called layers. The primary design objective is to design the layering architecture and associated signal processing so that the receiver can efficiently separate the individual layers from one another and can decode each of the layers effectively. The layering architectures in [5] were, however, inspired by signal processing considerations.

In [4], we generalized the notion of space-time layering independent of the signal processing employed at the receiver. In our framework, a *layer* in an $n \times T$ transmission resource array is identified by an indexing set $L \subset I_n \times I_T$ where the t-th symbol interval on antenna a belongs to the layer if and only if $(a,t) \in L$. This indexing set must satisfy the requirement that if $(a,t) \in L$ and $(a',t') \in L$, then either $t \neq t'$ or $a = a'$ (*i.e.,* that a is a function of t). This property ensures that each symbol interval within a layer is allocated to at most one antenna and

[1]This refers to the fact the information streams assigned to the different layers are independently encoded.

hence, all spatial interference experienced by the layer comes from outside the layer.

Consider a composite channel encoder γ consisting of n_1 constituent encoders $\gamma_1, \gamma_2, \ldots, \gamma_{n_1}$ operating on independent information streams. Let $\gamma_i : \mathcal{Y}^{k_i} \rightarrow \mathcal{Y}^{N_i}$, so that $k = k_1 + k_2 + \cdots + k_{n_1}$ and $N = N_1 + N_2 + \cdots + N_{n_1}$. Then, there is a partitioning $\underline{u} = \underline{u}_1 | \underline{u}_2 | \cdots | \underline{u}_{n_1}$ of the information vector $\underline{u} \in \mathcal{Y}^k$ into a set of disjoint component vectors $\underline{u}_i$, of length k_i and a corresponding partitioning $\gamma(\underline{u}) = \gamma_1(\underline{u}_1) | \gamma_2(\underline{u}_2) | \cdots | \gamma_{n_1}(\underline{u}_{n_1})$ of the code word $\gamma(\underline{u})$ into a set of constituent code words $\gamma_i(\underline{u}_i)$ of length N_i. In the layering approach, the space-time transmitter assigns each of the constituent code words $\gamma_i(\underline{u}_i)$ to one of the set of n_1 disjoint layers where, in general, $n_1 \leq n$.

The main innovation of the space-time layering approach is that it reduces the receiver design to the problem of joint multiuser detection and single user decoding. This problem has been extensively studied and several efficient algorithms with realizable receiver complexity have been proposed in the literature.

Because the layering approach imposes certain limitations on the transmitted signals (i.e., each layer is encoded independently), it faces the possibility of incurring a loss in capacity. The following proposition argues that, by carefully choosing the layering approach and the associated algorithms, one can still achieves the autocapacity with a layered space-time system.

Proposition 1: There exists a layered space-time system that achieves the autocapacity with a polynomial complexity receiver.

Proof: We will prove Proposition 1 by construction. First, we assume that the channel channel state information (CSI) is available at the receiver. Let $n = n_1 n_2$, where n_1 is the number of layers. Following in the footsteps of [2], we assume that $T = \beta n = \beta n_2 n_1$. Consider the layering assignment

$$L_k = \left\{ \left(\left\lfloor \frac{t}{\beta n_1} \right\rfloor n_1 + k, t \right) : 0 \leq t < T \right\}, \qquad \text{for } 1 \leq k \leq n_1 \tag{3}$$

shown pictorially in Fig. 2. It is clear from the figure that this layering assumption induces a block diagonal transmission format similar to that assumed in [2]. Define

$$H_i = \left[\underline{S}_{(i-1)(\beta n_1)+1}, \ldots, \underline{S}_{i\beta n_1} \right], \qquad \text{for } 1 \leq i \leq n_2 \tag{4}$$

where $\underline{S}_j$ is the jth column of the complex channel matrix S. Now, let's consider the ith diagonal block (i.e., $(i-1)\beta n_1 + 1 \leq t \leq i\beta n_1$). By letting $n_1 \rightarrow \infty$, the Shannon capacity of this block is

$$C_i = \lim_{n_1 \rightarrow \infty} \log_2 \left(\det \left(I_m + \frac{E_s}{N_0} H_i H_i^H \right) \right) \text{ b/channel use} \tag{5}$$

where $()^H$ denotes the hermitian operator. This capacity is achieved by using independent Gaussian code books for the different layers in that block. Now, averaging the capacity across the n_2 blocks and letting $n_2 \rightarrow \infty$, we obtain a lower bound on the capacity *possible* with for layered space-time systems

$$\begin{aligned} C_L \geq & \lim_{n_2 \rightarrow \infty} \frac{1}{n_2} \sum_{i=1}^{n_2} C_i \\ =& E[C_i] \text{ a.s.} \\ =& \lim_{n_1 \rightarrow \infty} E \left[\log_2 \left(\det \left(I_m + \frac{E_s}{N_0} H_i H_i^H \right) \right) \right] \\ =& m \log_2 (1 + \text{SNR}) = C_a \text{ b/channel use} \end{aligned} \tag{6}$$

where the second equality follows from the strong law of large numbers and the fact that the matrices $\{H_i\}$ are independent and identically distributed, the third equality follows from the linearity of the limit and expectation operators and the fourth equality follows from the autocapacity result in [2]. Since the lower bound is equal to the autocapacity, the proof is complete for the known CSI case. In addition, Varanasi and Guess showed that this capacity can be achieved with a polynomial complexity decision feedback MMSE receiver with asymmetric rate allocation to the different layers [9]. The extension to the case with no channel state information at the receiver follows the same argument as in [2]. □

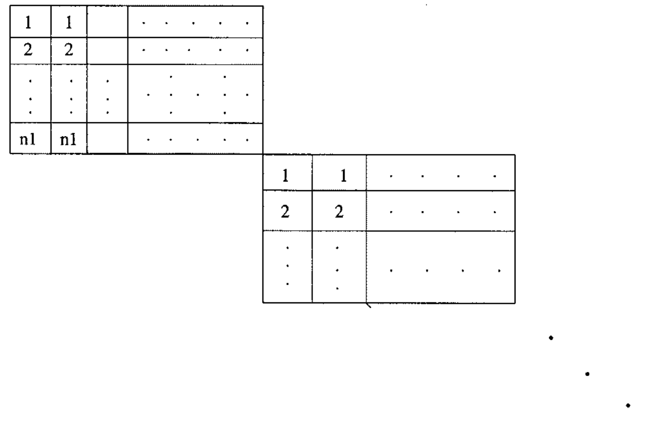


Fig. 2. A space-time layered system for block diagonal transmission.

The previous result argues that, for known CSI, the autocapacity can be achieved with a layered system using *independent* variable rate Gaussian code books in the different layers and MMSE decision feedback multiuser detection along with *single dimensional* successive decoding at the receiver. For the noncoherent scenario, a training based approach can be used for channel estimation. The loss in throughput incurred by the training interval vanishes as $T \rightarrow \infty$. It is clear that this receiver avoids the exponential complexity[2] involved in the previously proposed approach based on random unitary space-time signals [2]. This result, however, is limited by the asymptotic assumptions of a very large number of transmit antennas, a very long coherence interval, a highly incremental rate allocation strategy, and error free interference cancellation.

[2]The complexity of the receiver is only a polynomial function of the number of layers and the throughput is a linear function of the number of layers.

IV. THREADED SPACE-TIME LAYERING FOR AUTOCODING

Space-time systems optimized *only* to achieve the autocapacity *asymptotically* may not yield good performance in practical systems with limited numbers of transmit antennas and/or small coherence intervals. For example, in the layered architecture used to prove Proposition 1, each layer only spans n_2 antennas out of a total number of n antennas. This limitation, however, does not impose a loss in performance in the asymptotic scenario with very large number of transmit antennas, as shown earlier. In practice, the number of transmit antennas is limited and hence, it is critical for the layering approach to fully exploit the diversity offered by **all** transmit antennas. In an attempt to benefit from the autocoding phenomenon in practice, we proposed to use threaded space-time systems [4] with generalized bit interleaved space-time coded modulation, iterative signal processing and pilot symbol assisted channel estimation in the present scenario. In Section IV-A and Section IV-B, we assume the availability of channel state information at the receiver. Pilot symbol assisted channel estimation for the noncoherent scenario is briefly discussed in Section IV-C. The efficacy of this approach is supported by the simulation results reported in Section V.

The threaded layering approach efficiently exploits the diversity available in the MIMO channel by optimizing the encoding, interleaving and distribution of each layer's symbols among different antennas to maximize diversity for a given transmission rate, assuming no interference from the other layers. In this approach, the transmitter has available a disjoint set of layers, $\mathcal{L}_{Th} = \{L_1, L_2, \ldots, L_{n_1}\}$ and transmits the composite code word $\gamma(\underline{u}) = \gamma_1(\underline{u}_1)|\gamma_2(\underline{u}_2)|\cdots|\gamma_n(\underline{u}_{n_1})$ by sending $\gamma_i(\underline{u}_i)$ in layer L_i. The number of layers $n_1 \leq n$ is equal to the number of antennas transmitting simultaneously at any point of time. It is desirable to maximize n_1 to allow for the maximum transmission rate. The choice of n_1, however, must consider the efficiency of the signal processing algorithm as discussed in more details in Sections IV-C and V.

The layer set $\mathcal{L}_{Th}$ is designed so that each layer is active during all of the available symbol transmission intervals (i.e., $N_i = Tb$) and, over time, uses each of the n antennas equally often. Thus, during each symbol transmission interval, the layers each transmit a symbol using a different antenna; and, in terms of antenna usage, all of the layers are equivalent. Each layer in the new architecture is referred to as a *thread* where a thread can be defined as a layer with full spatial span "n" and full temporal span "T". This leads to the threaded layering set $\mathcal{L}_{Th}$ given by

$$L_i = \{(\lfloor t+i-1 \rfloor_n + 1, t) : 0 \leq t < T\}, \quad \text{for } 1 \leq i \leq n_1. \tag{7}$$

An example of this layering set is shown in Fig. 3 for a system with four transmit antennas and four threads. It is clear that in this system each thread utilizes all the available transmit antennas, even though all antennas may not be active simultaneously at any point of time (if $n_1 < n$). For simplicity of implementation, we further limit ourselves to the case in which the constituent codes are all of the same rate and have the same code word length: $N_i = N/n_1$ and $k_i = k/n_1$ for all i.

1	4	3	2	1	4	3
2	1	4	3	2	1	4
3	2	1	4	3	2	1
4	3	2	1	4	3	2

Fig. 3. A threaded space-time system with four transmit antennas and four threads. Each box represents one symbol transmission from the corresponding antenna. The numbers refer to the threads to which the boxes are assigned.

Under the error free interference cancellation assumption, threaded space-time layering allows for the design of component space-time codes that achieve the optimum tradeoff between diversity advantage and transmission rate for arbitrary numbers of transmit antennas and constellation sizes, as shown next.

A. Generalized Bit Interleaved Coded Modulation

In [10], Zehavi showed that the code diversity in fast Rayleigh fading channels can be improved by bitwise random interleaving prior to modulation. This approach results in a code diversity equal to the smallest number of distinct bits (rather than channel symbols) along any error event. In [11], the theory behind this technique was further developed and the term "bit interleaved coded modulation (BICM)" was coined for this approach. One of the advantages of BICM is that it offers *a clean separation* between the coding and modulation stages and hence, allows for more flexibility in changing the transmission rate while using the same binary channel code. This property motivates the consideration of this approach to construct *universal* threaded space-time component codes that maximize the diversity advantage for arbitrary numbers of transmit antennas and transmission rates.

In the absence of spatial interference from other threads, the channel experienced by an arbitrary thread L_i is equivalent to a single-input multi-output frequency nonselective block fading channel with n fading blocks per code word. The fading coefficients are constant across the fading block but change independently from block to block. It is easy to see that the bit-wise random interleaving used in the BICM is only optimized for the fast fading scenario and hence, may result in a loss in the diversity advantage in the current scenario. To avoid this loss, we generalize the original BICM construction by replacing the random interleaver with a periodic multiplexer followed by n random interleavers as shown in Fig. 4. The periodic multiplexer is used to distribute the encoded stream among the n fading blocks (i.e., transmit antennas), whereas the random bit interleaver used for each substream plays a double role. First, it distributes the encoded bits uniformly across the constellation positions, similar to the random intereleaver used in the original BICM construction [11]. Second, it allows for better convergence characteristics of the iterative receiver as discussed in [4], [12], [13]. The symbols are then

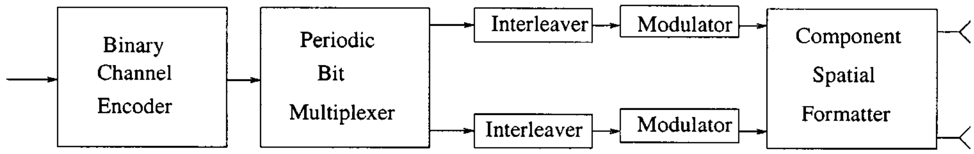


Fig. 4. Generalized bit interleaved coded modulation as a component space-time code in threaded layering. The component spatial formatter sends the modulated symbols in the time slots assigned to the thread.

sent in the time slots assigned to L_i[3] . In the generalized BICM, the channel encoder and periodic multiplexer should be jointly optimized to maximize the diversity advantage and it is straightforward to see that the diversity advantage is independent of the random interleavers and constellation mapping (i.e., modulator) modules. This property allows for using *universal* BICM binary codes with arbitrary constellations, where the size of the constellation is determined according to the desired throughput.

In the following, we develop sufficient conditions for binary codes and periodic multiplexers that achieve the maximum diversity advantage. Let C be the binary convolutional code of rate $1/n'$ used in Fig. 4. The encoder processes one binary input sequence $x(t)$ and produces n' coded output sequences $y_1(t), y_2(t), \ldots, y_{n'}(t)$ which are multiplexed together to form the output code word. A sequence $\{x(t)\}_{t=0}^{\infty}$ is often represented by the formal series $X(D) = x(0) + x(1)D + x(2)D^2 + \cdots$. We refer to $\{x(t)\} \leftrightarrow X(D)$ as a D-transform pair.

The action of the binary convolutional encoder is linear and is characterized by the so-called impulse responses $g_j(t) \leftrightarrow G_j(D)$ associating output $y_j(t)$ with the input $x(t)$.

The coded bits are to be distributed among n transmit antennas. For simplicity, we consider the case in which $s = n/n'$ is an integer and the coded bits are assigned to the antennas periodically (i.e., periodic bit multiplexing). Thus, for each of the coded bit streams $Y_i(D) \leftrightarrow \{y_i(t)\}$, the subsequence $y_i(0), y_i(s), y_i(2s), \ldots$ is assigned to antenna i; the subsequence $y_i(1), y_i(s+1), y_i(2s+1), \ldots$ is assigned to antenna $i + n'$; and so on. Alternate assignments such as symbol-based demultiplexing would also be possible and can be analyzed using the same framework.

In general, we partition the series $X(D)$ corresponding to $\{x(t)\}$ into its modulo s components $X_j(D)$ corresponding to the subsequences $\{x(st+j)\}_{t=0}^{\infty}$, where $j = 0, 1, 2, \ldots, s-1$. Then

$$X(D) = X_0(D^s) + D \cdot X_1(D^s) + \cdots + D^{s-1} \cdot X_{s-1}(D^s).$$

Similarly, we partition $G_i(D)$ into components $G_{i,j}(D)$ and $Y_i(D)$ into components $Y_{i,j}(D)$. The space-time code $\mathcal{C}$ under consideration therefore consists of the binary code C together with a spatial modulator function in which $Y_{i,j}(D)$ is assigned to antenna $n'j + i$.

[3]In the figure, the component spatial formatter sends the symbols in the time slots assigned to the thread.

By multiplying the expansions for $X(D)$ and $G_i(D)$ and collecting terms, one may show that the coded bit stream assigned to antenna $n'j + i$ is given by

$$Y_{i,j}(D) = \sum_{k=0}^{s-1} X_k(D) F_{n'j+i,k}(D)$$

where

$$F_{n'j+i,k}(D) = G_{i,j-k}(D) + D \cdot G_{i,j-k+s}(D).$$

In matrix form, we have

$$Y_{i,j}(D) = \mathbf{X}(D)\mathbf{F}_{n'j+i}(D)$$

which is the dot product of row vector $\mathbf{X}(D) = [X_0(D) \quad X_1(D) \quad \cdots \quad X_{s-1}(D)]$ and column vector

$$\mathbf{F}_{n'j+i}(D) = \begin{bmatrix} F_{n'j+i,0}(D) \\ F_{n'j+i,1}(D) \\ \vdots \\ F_{n'j+i,s-1}(D) \end{bmatrix}.$$

Now, the algebraic analysis technique proposed in [4] considers the rank of matrices formed by concatenating the column vectors $\{\mathbf{F}_\ell(D)\}$. Specifically, for $a_1, a_2, \ldots, a_n \in \mathbb{F}$, let

$$\mathbf{F}(a_1, a_2, \ldots, a_n) = [a_1\mathbf{F}_1 \quad a_2\mathbf{F}_2 \quad \cdots \quad a_n\mathbf{F}_n]$$

then [4, Theorem 6] applies directly: the spatial transmit diversity achieved by $\mathcal{C}$ is given by $d = n - v + 1$, where v is the smallest integer having the property that, whenever $a_0 + a_1 + \cdots + a_{n-1} = v$, the $s \times n$ matrix $\mathbf{F}(a_0, a_1, \ldots, a_{n-1})$ has full rank s. In particular, we note that the best possible spatial transmit diversity is $d = n - s + 1$. When $n' = n$, we have $s = 1$ so that full spatial transmit diversity $d = n$ is possible as expected. To illustrate the technique, we borrow the following simple example from [4]

Example: Consider the 4-state convolutional code with optimal $d_{\text{free}} = 5$ and generators $G_0(D) = 1 + D^2$ and $G_1(D) = 1 + D + D^2$. In the case of two transmit antennas, it is clear that the natural threaded space-time code achieves $d = 2$ level diversity.

In the case of four transmit antennas, we note that the rate 1/2 code can be written as a rate 2/4 convolutional code with generator matrix:

$$\mathbf{G}(D) = \begin{bmatrix} 1+D & 0 & 1+D & 1 \\ 0 & 1+D & D & 1+D \end{bmatrix}.$$

By inspection, every pair of columns is linearly independent. Hence, the natural periodic distribution of the code across four transmit antennas produces a threaded space-time code achieving the maximum $d = 3$ transmit spatial diversity.

For six transmit antennas, we express the code as a rate 3/6 code with generator matrix:

$$\mathbf{G}(D) = \begin{bmatrix} 1 & 0 & 1 & 1 & 1 & 1 \\ D & 1 & 0 & D & 1 & 1 \\ 0 & D & 1 & D & D & 1 \end{bmatrix}.$$

Every set of three columns in the generator matrix has full rank, so the natural space-time code achieves maximum $d = 4$ transmit diversity.

The same technique can be used with only minor modifications to analyze the performance with other multiplexing strategies (e.g., symbol based multiplexing [14]). Finally, we note that space-time BICM approaches were investigated independently in [15]–[17]. The generalized BICM proposed here is, however, *structurally* different from the previous works. The generalized BICM is coupled with the threaded layering design to maximize the diversity advantage in the absence of spatial interference. This maximization is achieved through a joint optimization of the channel encoder and periodic bit multiplexer. The other space-time BICM designs were based on a *random* distribution of the encoded bits across the different transmit antennas. This random distribution is more suited for fast fading channels [11] and may result in a loss in the diversity advantage in the scenario under consideration.

B. Iterative Signal Processing

The conditions developed in the previous section allow for designing component space-time codes with maximum spatial diversity. The main limitation, however, is the assumption that a genie were to cancel the other threads' interference at the receiver. Ultimately, the performance will hinge upon the efficiency of the signal processing algorithm in separating the signals from different threads. In Section III, we have argued that MMSE decision feedback multiuser detection with successive decoding can achieve the autocapacity. This design approach, however, entails utilizing *capacity achieving* channel codes to allow for error free interference cancellation. In practice, the decision feedback interference cancellation may suffer from error propagations. This receiver design also assumes a *highly incremental* rate allocation strategy which may be difficult to implement. The Turbo processing principle [18] can be efficiently used to avoid these drawbacks, as shown in [4]. In the Turbo receiver, shown in Fig. 5, a soft-input/soft-output (SISO) multidetector module based on the MMSE principle provides soft-decision estimates for the n_1 streams of data. The detected streams are decoded by the separate SISO binary channel decoders associated with the component channel codes. After each decoding iteration, the soft outputs from the channel decoders are used to refine the processing performed by the SISO multi detector. This architecture assumes symmetrical rate allocation and attempts to minimize the probability of error propagations through the use of SISO modules. It is also worth noting that, in the iterative receiver, each of the streams must be independently interleaved to facilitate convergence [12], [13]. We explicitly allowed for this independent interleaving option in our generalized BICM construction.

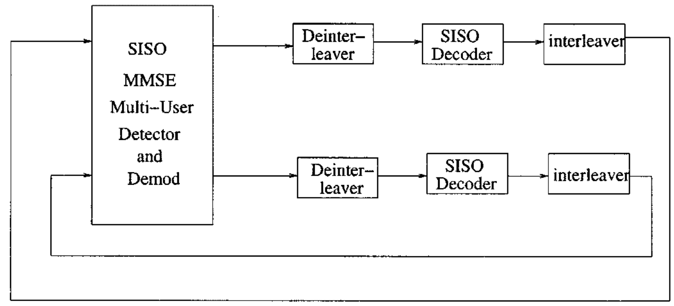


Fig. 5. A block diagram for the Turbo receiver. The number of SISO decoders is equal to the number of threads (i.e., each SISO decoder is assigned to one thread).

Before we present the iterative receiver for threaded systems with generalized BICM signals and arbitrary QAM constellations, we have the following result regarding symmetrical[4] layered systems with linear receivers

Proposition 2: The autocapacity of symmetrical layered space-time systems with linear receivers grows linearly with the number of receive antennas.

Proof: Please refer to the Appendix. □

The previous result shows that with symmetrical rate assignment and linear receivers, the capacity still grows linearly with the number of receiver antennas. The lower bound developed in the proof, however, predicts a significant loss in capacity, relative to C_a, especially with large SNRs. In the following, we briefly present the Turbo MMSE receiver for threaded systems that attempts to avoid this loss by iteratively cancelling the soft decoders' outputs.

The iterative MMSE receiver is adapted from our work [12] on iterative MMSE multiuser detectors for CDMA systems (Such receivers for CDMA applications were also investigated independently by Wang and Poor [13]). These works assumed binary phase-shift keying (BPSK) modulation for simplicity of presentation. In the following, we will outline the necessary modifications for the generalized BICM scenario with arbitrary constellations.

Let $y^{(i)}$ be the estimate of the ith antenna symbol at time t (the subscript t will be omitted for convenience) given by

$$y^{(i)} = \underline{w}_f^{(i)^T} \underline{r} + w_b^{(i)} \tag{8}$$

where $w_f^{(i)}$ is the $m \times 1$ optimized feedforward coefficients vector and $w_b^{(i)}$ is a single coefficient that represents the soft cancellation part. The coefficients $\underline{w}_f^{(i)}$, $w_b^{(i)}$ are obtained through minimizing the conditional mean square value of the error between the data symbol and its estimate (i.e., with the priors obtained from the decoders' extrinsic information). Now, let $\underline{S}^{(i)}$ be the $m \times 1$ complex channel vector of the ith transmit antenna; $S^{(n_1 \backslash i)}$ be the $m \times (n_1 - 1)$ matrix composed of the complex channel vectors of the other $n_1 - 1$ active transmit antennas at time t; $\underline{c}^{(n_1 \backslash i)}$ be the $(n_1 - 1) \times 1$ transmitted data vector form the other $n_1 - 1$ active transmit antennas. Using

[4]Contrary to the system used in the proof of the Proposition 1, now we focus on symmetrical systems where all the layers are identical.

standard minimization techniques, it is easily shown that the MMSE solutions for $\underline{w}_f^{(i)}$ and $w_b^{(i)}$ are given by[5]

$$\underline{w}_f^{(i)^T} = \underline{S}^{(i)^H} \left(A + B + R_n - FF^H\right)^{-1} \tag{9}$$

$$w_b^{(i)} = -\underline{w}_f^{(i)^T} F \tag{10}$$

where

$$A = \underline{S}^{(i)} \underline{S}^{(i)^H}; \tag{11}$$

$$B = S^{(n_1 \backslash i)} E\left[\underline{c}^{(n_1 \backslash i)} \underline{c}^{(n_1 \backslash i)^T}\right] S^{(n_1 \backslash i)^H} \tag{12}$$

$$F = S^{(n_1 \backslash i)} E\left[\underline{c}^{(n_1 \backslash i)}\right] \tag{13}$$

$$R_n = N_0 I_{m \times m} \tag{14}$$

$I_{m \times m}$ is the identity matrix of order m. After each decoding iteration, the soft outputs from the decoders are used to compute better *approximations* for the estimates used in (12) and (13). The independent random substream interleaving at the transmitter allows for the following approximation

$$E\left[c_t^{(i)} c_t^{(j)^*}\right] \approx E\left[c_t^{(i)}\right] E\left[c_t^{(j)}\right], \qquad \text{for } i \neq j, \tag{15}$$

which reduces the problem to finding $\forall i\ E[c_t^{(i)}]$ and $E[|c_t^{(i)}|^2]$. Computing these values from the *binary extrinsic information* depends on the modulation mapping μ. For general mapping operators, we have

$$E\left[c_t^{(i)}\right] = \sum_{d_{1t}^{(i)}, \ldots, d_{bt}^{(i)} \in \{-1,1\}} P\left(d_{1t}^{(i)}, \ldots, d_{bt}^{(i)}\right) \mu\left(d_{1t}^{(i)}, \ldots, d_{bt}^{(i)}\right) \tag{16}$$

$$E\left[|c_t^{(i)}|^2\right] = \sum_{d_{1t}^{(i)}, \ldots, d_{bt}^{(i)} \in \{-1,1\}} P\left(d_{1t}^{(i)}, \ldots, d_{bt}^{(i)}\right) \left|\mu\left(d_{1t}^{(i)}, \ldots, d_{bt}^{(i)}\right)\right|^2 \tag{17}$$

where $d_{1t}^{(i)}, \ldots, d_{bt}^{(i)}$ are the binary coordinates for the symbol $c_t^{(i)}$. The independent sub-stream interleaving also allows for the following approximation:

$$P\left(d_{1t}^{(i)}, \ldots, d_{bt}^{(i)}\right) \approx \prod_{k=1}^{b} P\left(d_{kt}^{(i)}\right) \tag{18}$$

where the marginal probability can now be obtained from the extrinsic information as

$$P\left(d_{kt}^{(i)} = 1\right) = 1 - P\left(d_{kt}^{(i)} = -1\right) = \frac{e^{\lambda_{kt}^{(i)}}}{1 + e^{\lambda_{kt}^{(i)}}} \tag{19}$$

where $\lambda_{kt}^{(i)}$ is the extrinsic information corresponding to the binary symbol $d_{kt}^{(i)}$ [19]. Note that in the first iteration, one assumes uniform distribution for all binary symbols. It is clear that (16) and (17) entail an exponentially growing complexity with the number of transmitted bits per symbol (i.e., transmission rate). In the special case of square QAM constellations with set partition mapping, b is even and the baseband modulation mapping is given by

$$\mu(d_1, \ldots, d_b) = \frac{1}{Z} \sum_{k=1}^{b/2} d_k 2^{k-1} + j d_{(b/2+k)} 2^{k-1} \tag{20}$$

where Z is a normalization constant; $j = \sqrt{-1}$. This observation allows for the following linear implementations of (16) and (17):

$$E\left[c_t^{(i)}\right] = \frac{1}{Z} \sum_{k=1}^{b/2} E\left[d_{kt}^{(i)}\right] 2^{k-1} + jE\left[d_{(b/2+k)t}^{(i)}\right] 2^{k-1} \tag{21}$$

$$E\left[\left|c_t^{(i)}\right|^2\right] = \left|E\left[c_t^{(i)}\right]\right|^2 + \frac{1}{Z^2} \sum_{k=1}^{b/2} 2^{2(k-1)} \cdot \left(2 - \left|E\left[d_{kt}^{(i)}\right]\right|^2 - \left|E\left[d_{(k+b/2)t}^{(i)}\right]\right|^2\right) \tag{22}$$

where $E[d_{kt}^{(i)}]$ can be easily computed using (19). The binary log-likelihood ratios necessary for the next decoding iteration can now be computed from $y^{(i)}$ using one of the two approaches presented in [10] and [11].

C. Parameter Optimization and Channel Estimation

In the coherent scenario, when the CSI is available *a priori* at the receiver, the transmission rate of the threaded system is given by

$$R_c = n_1 b r \text{ b/channel use} \tag{23}$$

where r is the binary channel code rate. From (23), it is clear that to maximize the transmission rate for a fixed constellation size, it is desirable to maximize the number of threads n_1 and coding rate r. Increasing the coding rate, however, will decrease the maximum possible transmit diversity, as shown by the fundamental limit $(d_{\max} = \lfloor n(1-r) \rfloor + 1)$ [4], [14]. This loss in diversity order may entail a loss in the system power efficiency which will depend on the operating SNR. Similarly, increasing n_1 beyond a certain value may limit the ability of the iterative decoder to converge and hence, entails a significant loss in the system power efficiency. These two reasons suggest that there exist optimum values for n_1, b and r that determine the maximum transmission rate for the operating SNR and bit error rate requirement. Unfortunately, at the time, we don't have an analytical approach for optimizing these values and hence, we will resort to empirical techniques in the numerical results section.

In the noncoherent scenario, orthogonal pilot symbols are inserted in the transmitted stream to facilitate channel estimation at the receiver [20]. The transmission rate is now given by

$$R_{nc} = \frac{T - \gamma n}{T} n_1 b r \ \text{ b/channel use} \tag{24}$$

where T is the coherence interval and $\gamma \geq 1$. It is clear from (24) that choosing γ should be based on the tradeoff between transmission rate and channel estimation errors. In the absence of analytical tools to find optimal values for γ in the proposed system, we use the result in [21] which argues that for systems with small coherence intervals and large SNRs, γ should be set to one. In fact, for systems with very small coherence intervals,

[5]The expectations in these expressions refer, more rigorously, to the conditional expectations knowing the extrinsic information from the previous decoding iteration.

it may be beneficial not to use all the transmit antennas to avoid the throughput loss necessary to estimate all the corresponding channels. This contrasts the coherent scenario where all the antennas should be utilized, although not necessarily at the same point of time. This approach would modify (24) to

$$R_{nc} = \frac{T - \gamma n_u}{T} n_1 br \text{ b/channel use} \quad (25)$$

where n_u is the number of used transmit antennas. In this case, the threaded approach will be modified such that each thread only spans n_u antennas. It is worth noting that the number of simultaneous transmissions remains n_1 where, in general, $n_1 \leq n_u \leq n$.

As a final remark, we note that the Turbo processing approach [18] can be also used to enhance the channel estimation process. This approach, however, entails additional complexity and will not be pursued further in this paper.

V. Simulation Results

In this section, we report simulation results that demonstrate the efficacy of threaded space-time systems in realizing significant fractions of the autocapcacity with limited numbers of transmit antennas and small coherence intervals. To limit the receiver complexity, we use the 4-state $(5_8, 7_8)$ convolutional code as the component code in the generalized BICM scheme. The soft output Viterbi algorithm is used for the SISO decoders and the number of iterations is ten. The SNRs in the figures correspond to the total energy transmitted from all antennas. In all the simulation results, the interleaver length corresponds to one coherence interval.

A. Coherent Scenario

Here, we assume that channel state information is available *a priori* at the receiver. Fig. 6 reports the performance of the threaded layering approach for systems with eight receive antennas and different numbers of transmit antennas. We assume that all transmit antennas will be active simultaneously (i.e., $n_1 = n$). The results in the figure correspond to QPSK modulation with rate 1/2 coding, so the overall transmission rate is equal to the number of transmit antennas. To compare the proposed technique with the autocapacity, we choose a target bit error rate of 10^{-5}. The efficiency of the system is therefore defined as

$$\eta_c = \frac{R_c}{C_a\left(\mathrm{SNR}_t\right)} \quad (26)$$

where SNR_t is the SNR corresponding to the 10^{-5} bit error rate and $C_a(\mathrm{SNR}_t)$ is the autocapacity at this particular SNR. Fig. 7 reports the efficiency with different numbers of transmit antennas, where it is shown that the maximum efficiency is achieved with ten transmit antennas[6]. This result supports our earlier claim that, contrary to the optimal receiver scenario, there is an optimal **finite** value for the number of **simultaneous** transmissions, i.e., number of threads, when the Turbo receiver is used. The same comparison of Fig. 7 is repeated for the cases of six and four receive antennas in Figs. 8 and 9, respectively.

[6]It is worth noting that the results in [3] indicate efficiencies in the range of 20%–25%, however, for the noncoherent scenario.

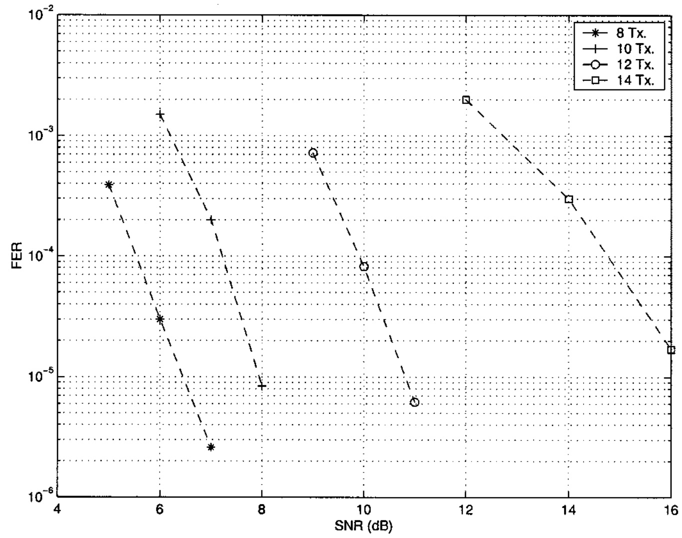


Fig. 6. Performance of threaded space-time layering with eight receive antennas and QPSK modulation.

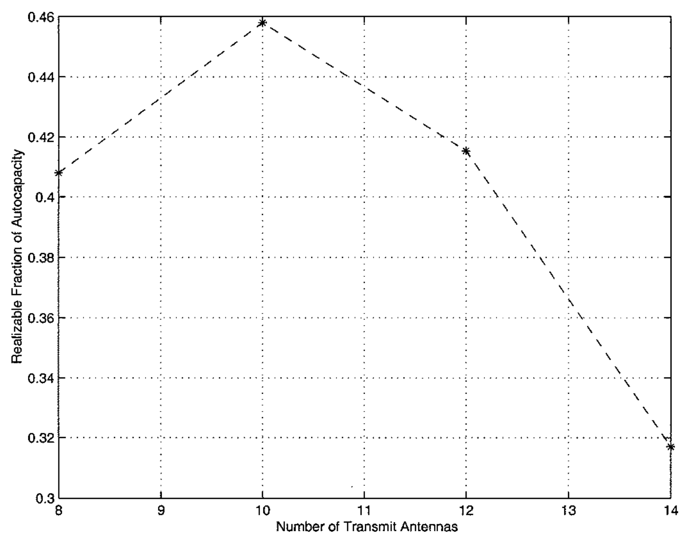


Fig. 7. Efficiency of threaded space-time layering with eight receive antennas, QPSK modulation and different numbers of transmit antennas.

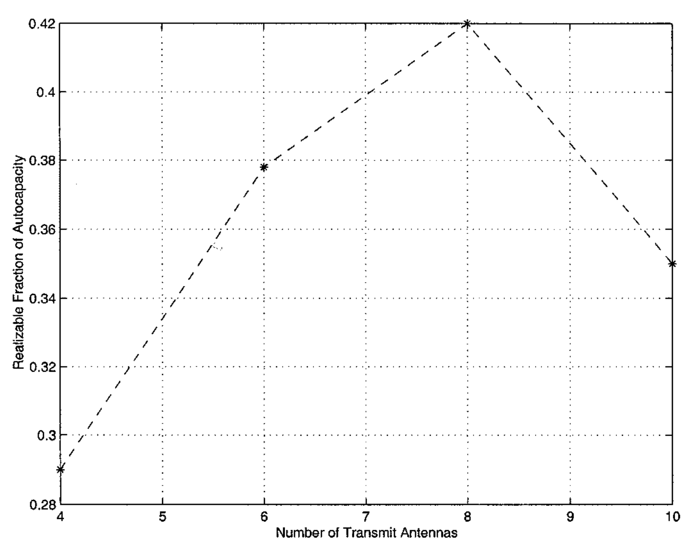


Fig. 8. Efficiency of threaded space-time layering with six receive antennas, QPSK modulation and different numbers of transmit antennas.

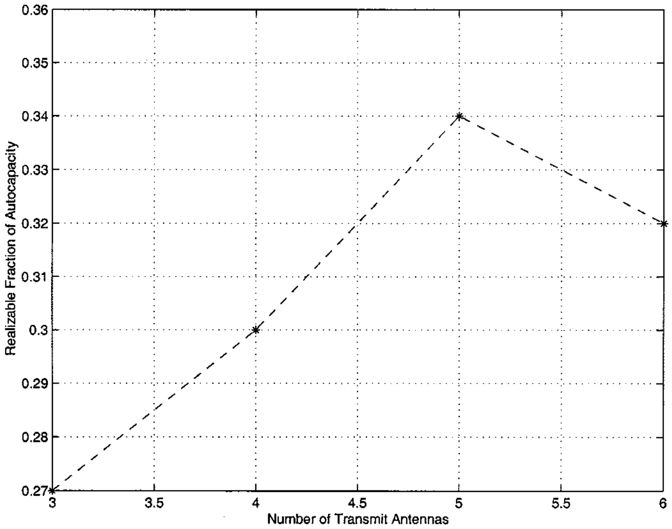


Fig. 9. Efficiency of threaded space-time layering with four receive antennas, QPSK modulation and different numbers of transmit antennas.

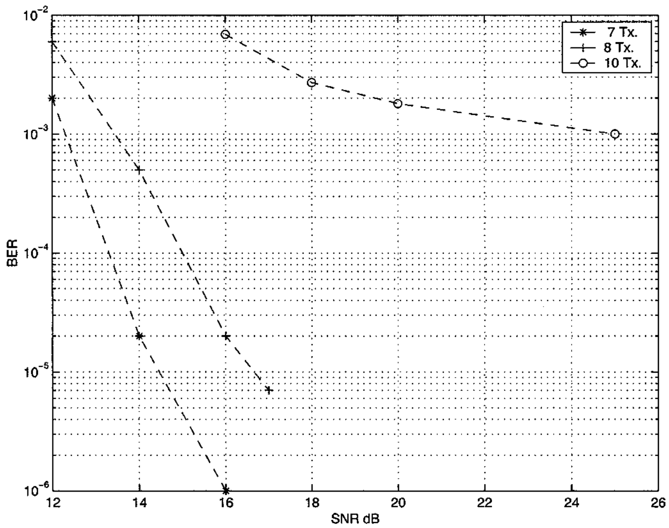


Fig. 10. Performance of threaded space-time layering with eight receive antennas and 16–QAM modulation.

Comparing these results, one observes that the maximum efficiency increases with the number of receive antennas. One possible explanation for this observation is the better convergence characteristics of Turbo receivers in large systems [12].

In Figs. 10 and 11, we consider the case with 16–QAM modulation, rate 1/2 coding and eight receive antennas[7] . The transmission rate in this scenario is twice the number of transmit antennas. Similar to the QPSK case, it is shown that there is an optimum value for the number of transmit antennas[8] . The optimum value, however, is different from that in the QPSK case.

It is worth noting that the optimum value for transmit antennas reported in the figures corresponds, more accurately, to

[7]It is interesting to observe the error floor at 10^{-3} bit error rate for the ten transmit antenna case.

[8]We have reported an autocapacity equal to zero in the case of ten transmit antennas due to the error floor in Fig. 10 which does not allow for achieving $\mathrm{BER} = 10^{-5}$ for any input SNR

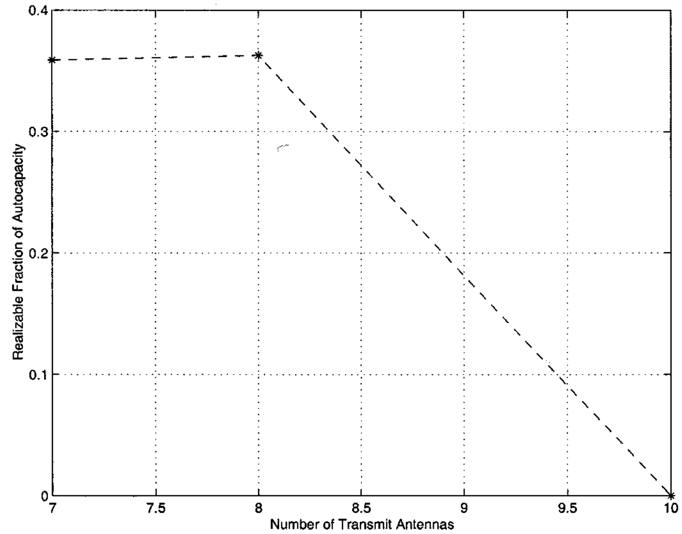


Fig. 11. Efficiency of threaded space-time layering with eight receive antennas, 16–QAM modulation and different numbers of transmit antennas.

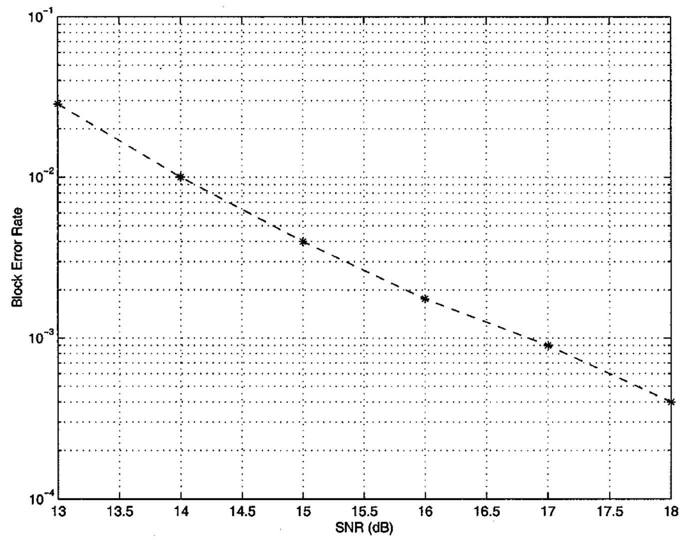


Fig. 12. Performance of threaded space-time layering in a noncoherent systems with four receive antennas, $(n, T) = (4, 20)$, rate 3/4 coding and QPSK modulation.

the number of threads "n_1" (in our case we assume the number of threads to be equal to the total number of transmit antennas). As outlined in the paper, we can always increase "n" without changing "n_1" to obtain higher diversity advantages. In the current scenario, this increase would not entail any *price* since we assume prior knowledge of the channel state information. In the *more realistic* noncoherent scenario, the price would be a loss in the transmission rate to allow for longer training sequences for channel estimation.

B. Non-Coherent Scenario

In this section, we compare the proposed threaded approach with some of the results in [3], [6]. In Fig. 12, we report the block probability of error[9] for the threaded approach with four

[9]The block probability of error was reported here to have a fair comparison with results in [3].

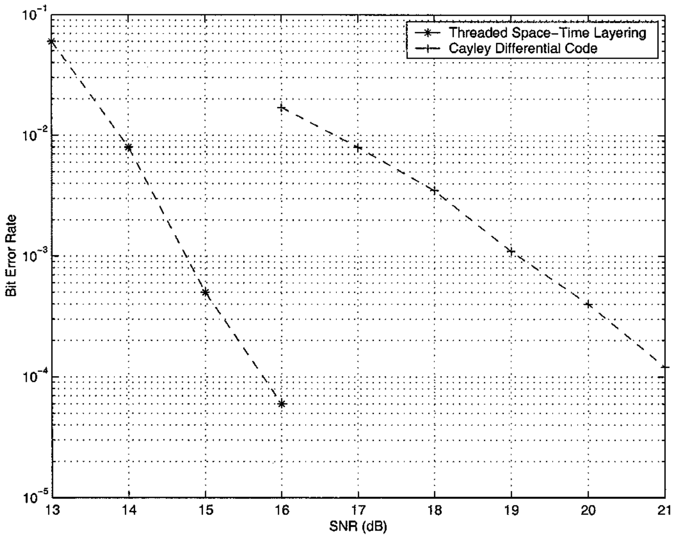


Fig. 13. Performance of threaded space-time layering and Cayley differential coding in a system with 8 transmit and 12 receive antennas with 16 bits/sec/Hz.

receive antennas and $(T,n) = (20,4)$. The coded stream is punctured to a obtain a rate 3/4 code and QPSK modulation is used. After accounting for the training sequence and the terminating bits for the convolutional code, the overall rate is 4.63 b/s/Hz which corresponds to approximately 20% of the the autocapacity. Comparing these results with [3], one observes that the proposed threaded system achieves the same performance at $\mathrm{SNR} = 18$ dB as the upper bound on the structured unitary approach with $(T,n) = (8,3)$[10] and the single rotation matrix approach with $(T,n) = (16,7)$[11] [3]. The main advantage of the threaded layering approach, however, is that it avoids the exponentially growing complexity exhibited in the other approaches [3] and still achieves a significant fraction of the autocapacity at low block error rates.[12]

Fig. 13 compares the performance of the threaded layering technique and the Cayley differential coding approach [6] in a system with eight transmit and twelve receive antennas. This comparison is motivated by the fact that both of these approaches avoid the exponentially growing complexity of the receiver in [3]. An accurate complexity comparison between the two approaches is, however, beyond the scope of this paper due to the structural differences between the two approaches. For the threaded approach, we use 16–QAM modulation and rate 3/4 coding. For the Cayley differential codes, the channel is assumed to be fixed for two consecutive *matrix* symbols (i.e., 16 transmission intervals) [6]. In the proposed case, however, we assume a coherence interval equal to 26 symbols[13]. As shown in the figure, the proposed threaded approach achieves a significant gain compared with the Cayley differential code, however, with a longer coherence interval. We also note that the gain of the threaded approach is limited to systems with reasonable numbers of receive antennas, so the Cayley differential coding approach remains the preferred solution in systems with limited numbers of receive antennas.

[10]See [3, Fig.3].

[11] See [3, Fig. 4].

[12]Although we did not extend our performance curve below block error rate of 10^{-4} due to the excessive simulation time needed, we still believe that this operating range would be satisfactory for many practical applications.

[13]This way, we achieve the same throughput as the Cayley code after accounting for the training sequence and terminating bits.

VI. Conclusions

In this paper, we considered the design of layered space-time systems for autocoding. We showed that carefully chosen layered space-time systems can achieve the autocapacity while avoiding the prohibitive complexity encountered in the unitary matrices signalling approach. We also addressed the different issues pertaining to practical systems with limited numbers of transmit antennas and small coherence intervals. For such systems, we investigated the design of threaded space-time layering with generalized BICM, iterative signal processing and pilot symbol assisted channel estimation as a method to exploit the autocoding phenomenon in practice. Finally, simulation results that exhibit the excellent performance of the threaded space-time approach in certain representative scenarios were presented.

Appendix
Proof for Proposition 2

First, we assume the availability of channel state information at the receiver. Let C_{LL} denote the autocapacity possible with layered systems and linear receivers. To prove our claim, it suffices to find a lower bound on C_{LL} that grows linearly with the number of receive antenna. To this end, consider the same layering assignment used in the proof of Proposition 1. Also, based on the same argument used in the proof of Proposition 1, we can see that

$$C_{LL} = E\left[C_{LL}^{(i)}\right] \tag{27}$$

where $C_{LL}^{(i)}$ is the capacity associated with the ith transmission block. To obtain a lower bound on $C_{LL}^{(i)}$, we consider a system that uses independent Gaussian code books with equal rates and powers in the different layers (i.e., by assuming a specific structure, we guarantee to obtain a lower bound). Due to the symmetry, we can now focus only on the first layer. The Gaussianity of the code books implies that, conditioned on the path gains, the interference at the jth receive antenna has a zero mean complex Gaussian distribution with variance

$$\sigma_I^2 = \frac{E_t}{n_1}\sum_{k=2}^{n_1}\left|\hat{\alpha}^{(kj)}\right|^2 \tag{28}$$

where E_t is the total energy transmitted from all the antennas; $\hat{\alpha}^{(kj)}$ is the path gain from the kth *active* transmit antenna to the jth receive antenna. Letting the number of active transmit antennas grow, we obtain

$$\begin{aligned}\sigma_I^2 &= \lim_{n_1\to\infty}\frac{E_t}{n_1}\sum_{k=2}^{n_1}\left|\hat{\alpha}^{(kj)}\right|^2\\ &= \frac{n_1-1}{n_1}E_t \text{ a.s.}\end{aligned} \tag{29}$$

which follows from the strong law of large number. It is clear that σ_I^2 is independent of the path gains which implies that the unconditional distribution of the interference will also be Gaussian in the limit. Also, the fact that σ_I^2 is independent of the receive antenna implies that the optimal linear detector is given by

$$y_t^{(1)} = \underline{w}^T \underline{r}_t = \left[\hat{\alpha}^{*(11)}, \ldots, \hat{\alpha}^{*(1m)}\right] \underline{r}_t \tag{30}$$

which maximizes the SNR (i.e., in the limit when $n_1 \rightarrow \infty$, the MMSE receiver is the same as the matched filter receiver). Using the symmetry and the Gaussian nature of the linear detector output, the capacity of the ith transmission block can be lower bounded by

$$C_{LL}^{(i)} \geq \lim_{n_1 \rightarrow \infty} n_1 \log_2 \left[1 + \frac{E_t \sum_{l=1}^{m} \left|\hat{\alpha}^{(1l)}\right|^2}{n_1 N_0 + (n_1 - 1) E_t}\right] = \log_2 e \frac{E_t}{E_t + N_0} \sum_{l=1}^{m} \left|\hat{\alpha}^{(1l)}\right|^2. \tag{31}$$

Now the lower bound on C_{LL} is readily available as

$$C_{LL} \geq \log_2 e \frac{E_t}{E_t + N_0} \sum_{l=1}^{m} E \left|\hat{\alpha}^{(1l)}\right|^2 = \log_2 e \frac{mE_t}{E_t + N_0}. \tag{32}$$

The extension to the noncoherent scenario follows the same argument as [2] with no loss in capacity as $T \rightarrow \infty$.

References

[1] G. J. Foschini and M. Gans, "On the limits of wireless communication in a fading environment when using multiple antennas," *Wireless Pers. Commun.*, vol. 6, pp. 311–335, Mar. 1998.

[2] B. Hochwald, T. Marzetta, and B. Hassibi, "Space-time autocoding," *IEEE Trans. Inform. Theory*, Nov. 2001.

[3] T. Marzetta, B. Hassibi, and B. Hochwald, "Structured unitary space-time autocoding constellations," *IEEE Trans. Inform. Theory*, vol. 48, pp. 942–950, Apr. 2000.

[4] H. El Gamal and A. R. Hammons Jr, "A new approach to layered space-time coding and signal processing," *IEEE Trans. Inform. Theory*, Sept. 2001.

[5] G. J. Foschini, "Layered space-time architecture for wireless communication in fading environments when using multiple antennas," *Bell Labs. Tech. J.*, vol. 2, pp. 2321–2334, Autumn 1996.

[6] B. Hochwald and B. Hassibi, "Cayley differential unitary space-time codes," *IEEE Trans. Inform. Theory*, vol. 48, pp. 1485–1503, 2002.

[7] J.-C. Guey, M. R. Bell, M. P. Fitz, and W.-Y. Kuo, "Signal design for transmitter diversity wireless communication systems over Rayleigh fading channels," in *IEEE Vehicular Technology Conf.*, Atlanta, GA, 1996, pp. 136–140.

[8] 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. 774–765, Mar. 1998.

[9] M. Varanasi and T. Guess, "Optimum decision feedback multi equalization and successive decoding achieves the total capacity of the Gaussian multiple access channel," in *Proc. Asilomar Conf. Signals, Systems and Computers*, 1997.

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

[11] G. Caire, G. Taricco, and E. Bigliere, "Bit interleaved coded modulation," *IEEE Trans. Inform. Theory*, vol. 44, pp. 927–946, May 1998.

[12] H. E. Gamal and E. Geraniotis, "Iterative multi detection for coded CDMA signals in AWGN and Rayleigh fading channels," *IEEE J. Select. Areas Commun.*, vol. 18, pp. 30–41, Jan. 2000.

[13] X. Wang and H. V. Poor, "Iterative (turbo) soft interference cancellation and decoding for coded CDMA," *IEEE Trans. Commun.*, vol. 47, pp. 46–64, July 1999.

[14] A. Lapidoth, "The performance of convolutional codes on the block erasure channel using various finite interleaving techniques," *IEEE Trans. Inform. Theory*, vol. 43, pp. 1459–1473, Sept. 1997.

[15] E. Biglieri, G. Taricco, and E. Viterbo, "Bit interleaved time-space codes for fading channels," presented at the Conf. on Information Sciences and Systems, Princeton, NJ, Mar. 2000.

[16] Z. Hong and B. L. Hughes, "Robust space-time trellis codes based on bit interleaved coded modulation," presented at the Conf. on Information Sciences and Systems, MD, Mar. 2001.

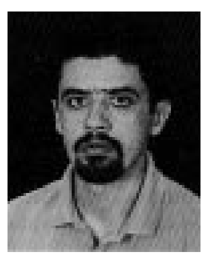
[17] J. J. Boutros, F. Boixadera, and C. Lamy, "Bit interleaved coded modulation for multi-input multi-output fading channels," presented at the IEEE 6th Int. Symp. on Spread Spectrum Techniques and Applications, NJ, Sept. 2000.

[18] J. Hagenauer, "The turbo principle: Tutorial introduction and state of the art," in *Int. Symp. on Turbo Codes and Related Topics*, Brest, France, Sept. 1997, pp. 1–9.

[19] "Telecom. and Data Acquisition Progress Rep.," Jet Propulsion Lab., Pasadena, CA, Nov. 1996.

[20] A. Naguib, V. Tarokh, N. Seshadri, and R. Calderbank, "A space-time coding modem for high data rate wireless communications," *IEEE J. Select. Areas Commun.*, vol. 16, pp. 1459–1478, Oct. 1998.

[21] B. Hassibi and B. Hochwald, "How much training is needed in multiple antenna wireless links," IEEE Trans. Inform. Theory, submitted for publication.



Heshman El Gamal (M'99) received the B.S. and M.S. degrees in electrical engineering from Cairo University, Cairo, Egypt, in 1993 and 1996, respectively, and the Ph.D. degree in electrical engineering from the University of Maryland at College Park, MD, in 1999.

From 1993 to 1996, he served as a Project Manager in the Middle East Regional Office of Alcatel Telecom. From 1996 to 1999, he was a Research Assistant in the Department of Electrical and Computer Engineering, the University of Maryland at College Park, MD. From February 1999 to January 2001, he was with the Advanced Development Group, Hughes Network Systems, Germantown, MD, as a Senior Member of the Technical Staff. In the Fall of 1999, He served as a lecturer at the university of Maryland at college park. Starting from January 2001, he assumed his current position as an Assistant Professor in the Electrical Engineering Department at the Ohio State University, Columbus. His research interests include spread spectrum communication systems design, multidetection techniques, coding for fading channels with emphasis on space-time codes, and the design and analysis of codes based on graphical models. He currently serves as an Associate Editor for the IEEE TRANSACTIONS ON COMMUNICATIONS.