We present in Fig. 5 the performance of $\mathbf{\Theta}_2$ (using a 4-QAM constellation), the Alamouti scheme (using a 16-QAM constellation), and the V-BLAST system with two transmit antennas (using a 4-QAM constellation) antennas at a spectral efficiency of 4 bits PCU and with $N = 2$ receive antennas. It is seen that $\mathbf{\Theta}_2$ has a gain of about 0.3 dB over the Alamouti scheme.

Fig. 6 shows the performance of the modulation $\mathbf{\Theta}_{3,2}$ (using a 4-QAM constellation), the DAST modulation Ξ_3 (using a 64-QAM constellation), and the Alamouti modulation (using a 64-QAM constellation), with $N = 3$ receive antennas at a spectral efficiency of 6 bits PCU. Modulation $\mathbf{\Theta}_{3,2}$ has a coding gain of about 7 dB over the Alamouti scheme, with the same order of total achieved diversity. We observe that for SNR of less than 20 dB, the ST modulation Ξ_3 has slightly worse performance than the Alamouti modulation, although it achieves higher diversity ($3N$ compared to $2N$ achieved by the Alamouti scheme). This is due to the better distribution of the product distance of the Alamouti modulation matrix over the SNR range considered.

VI. Conclusion

In this correspondence, we have studied the properties of some bandwidth-efficient ST modulations over M transmit antennas and T symbol periods with a rate of M symbols per channel use and a transmit diversity order of $\min(M, T)$ under quasi-static fading [1]. These modulations transmit rotated versions of MT-dimensional 4-QAM constellations. The extension of the proposed modulations to QAM constellations of higher spectral efficiency was considered. Under fast fading, we have shown that the proposed modulations can be generalized to any number of transmit antennas M and any number of symbol periods T, and still transmit at a rate of M symbols per channel use achieving a transmit diversity of T.

Acknowledgment

The authors are grateful to the anonymous reviewers and to the Associate Editor Giuseppe Caire for their helpful comments. M. O. Damen would like to thank Amin Shokrollahi for many helpful discussions on the quadratic form in (20).

References

[1] M. O. Damen, "Joint coding/decoding in a multiple access system, application to mobile communications," Ph.D. dissertation, ENST de Paris, Paris, France, Oct. 1999. [Online]. Available: http://www.ee.ualberta.ca/damen.

[2] E. Telatar, "Capacity of multi-antenna Gaussian channels," AT & T Bell Laboratories, Murray Hill, NJ, Tech. Rep., 1995.

[3] G. J. Foschini and M. J. Gans, "On limits of wireless communications in a fading environment when using multiple antennas," *Wireless Personal Commun.*, vol. 6, no. 3, pp. 311–335, 1998.

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

[5] V. Tarokh, N. Seshadri, and A. Calderbank, "Space-time codes for high data rate wireless communications: Performance criterion and code construction," *IEEE Trans. Inform. Theory*, vol. 44, pp. 744–765, Mar. 1998.

[6] G. D. Golden, G. J. Foschini, R. A. Valenzuela, and P. W. Wolniasky, "Detection algorithm and initial laboratory results using V-BLAST space-time communication architecture," *IEE Electron. Lett.*, vol. 35, pp. 14–16, Jan. 1999.

[7] M. O. Damen, A. Chkeif, and J.-C. Belfiore, "Lattice codes decoder for space-time codes," *IEEE Commun. Lett.*, vol. 4, pp. 161–163, May 2000.

[8] S. M. Alamouti, "A simple transmit diversity technique for wirelesscommunications," *IEEE J. Select. Areas Commun.*, vol. 16, pp. 1451–1458, Oct. 1998.

[9] V. Tarokh, H. Jafarkhani, and A. R. Calderbank, "Space-time block codes fromorthogonal designs," *IEEE Trans. Inform. Theory*, vol. 45, pp. 1456–1466, July 1999.

[10] M. O. Damen, K. Abed-Meraim, and J.-C. Belfiore, "Diagonal algebraic space-time block codes," *IEEE Trans. Inform. Theory*, vol. 48, pp. 628–636, Mar. 2002.

[11] B. Hassibi and B. Hochwald, "High-rate codes that are linear in space and time," *IEEE Trans. Inform. Theory*, vol. 48, pp. 1804–1824, July 2002.

[12] J. M. Wozencraft and I. M. Jacobs, *Principles of Communication Engineering*. New York: Wiley, 1965.

[13] X. Giraud, E. Boutillon, and J.-C. Belfiore, "Algebraic tools to build modulation schemes for fading channels," *IEEE Trans. Inform. Theory*, vol. 43, pp. 938–952, May 1997.

[14] J. Boutros and E. Viterbo, "Signal space diversity: A power and bandwidth efficient diversity technique for the Rayleigh fading channel," *IEEE Trans. Inform. Theory*, vol. 44, pp. 1453–1467, July 1998.

[15] M. O. Damen, K. Abed-Meraim, and J.-C. Belfiore, "Generalized sphere decoder for asymmetrical space-time communication architecture," *IEE Electron. Lett.*, vol. 36, p. 166, Jan. 2000.

[16] M. O. Damen, K. Abed-Meraim, and M. S. Lemdani, "Further results on the sphere decoder," in *Proc. IEEE Int. Symp. Information Theory*, Washington, D.C., June 2001, p. 333.

[17] J. G. Proakis, *Digital Communications*, 4th ed. New York: McGraw-Hill, 2000.

Differential Space–Time Turbo Codes

Christian Schlegel, *Senior Member, IEEE*, and
Alex Grant, *Senior Member, IEEE*

***Abstract*—Serial concatenation of simple error control codes and differential space–time modulation is considered. Decoding is performed iteratively by passing symbol-wise *a posteriori* probability values between the decoders of the inner space–time code and the outer code. An extrinsic information transfer analysis is used to predict thresholds for outer convolutional codes of various memory orders and a simple outer parity-check code. This parity-check code is well matched to the inner differential space–time code and achieves a bit-error rate (BER) of 10^{-6} less than 2 dB from the Shannon capacity of the fast fading multiple antenna channel. The differential space–time code can also be used to generate *a priori* information in the absence of channel knowledge. This information can be exploited by a channel estimator inserted into the decoding iteration. It is demonstrated that the inner space–time code provides soft training symbols from periodically inserted training symbols. The reliability of these soft training symbols does not depend on the speed of the channel variations, but on the structure of the inner code and the signal-to-noise ratio (SNR). Simulation studies confirm these findings and show that the proposed system with no initial channel knowledge achieves a performance very close to that of the system with perfect channel knowledge.**

***Index Terms*—Channel estimation, differential modulation, iterative decoding, space–time codes.**

I. Introduction

The tremendous growth of wireless communications services over the past decade motivates the design of power- and frequency-efficient

Manuscript received December 13, 2001; revised May 1, 2003. The work of C. Schlegel was supported in part by the National Science Foundation under Grants CCR-9732962 and ECS-0121389. The work of A. Grant was supported in part by ARC under Grants C00002232, X00106499, and DP0209658. The material in this correspondence was presented in part at the IEEE Information Theory Workshop, Cairns, Australia, September 2001.

C. Schlegel is with the Department of Electrical and Computer Engineering, University of Alberta, Edmonton, AB T6G 2V4, Canada.

A. Grant is with the Institute for Telecommunications Research, University of South Australia, Mawson Lakes, SA 5095, Australia.

Communicated by R. Koetter, Associate Editor for Coding Theory.
Digital Object Identifier 10.1109/TIT.2003.815818

portable communications devices. Applications such as mobile computing, video transmission to mobile devices, and other high-speed data services have increased the demand for higher data rates, from 64 bits/s in enhancements to current second-generation cellular systems, to 2 Mbit/s in next-generation mobile services, and beyond.

Regardless of advanced coding techniques such as turbo codes [1], channel capacity remains an unmovable barrier. The need for higher and higher data rates can no longer be supported by simply allocating wider frequency bands, and thus other methods of capacity increase have been researched heavily in the past decade. One promising technique is the use of multiple transmit and receive antennas. Such systems can theoretically increase capacity by up to a factor equaling the number of transmit and receive antennas in the array [2]–[6].

Spatial diversity has been proposed for support of very high rate data users within third-generation wide-band code-division multiple-access (CDMA) systems such as cdma2000 [7]. Using multiple antennas, these systems achieve gains in link quality and, therefore, capacity. Classically, diversity has been exploited through the use of either beam steering (for antenna arrays with correlated elements) or through diversity combining (for independent antenna arrays) [8], [9]. Use of these array processing techniques can achieve any combination of the following: a) reduction of multiple-access interference through the "nulling" of strong interferers. Such techniques are complementary to (and share the mathematical formulations of) multiple-user receivers such as the decorrelator and minimum mean-square error (MMSE) filter [10]; b) mitigation of fading effects by averaging over the spatial properties of the fading process. This is a dual of interleaving techniques which average over the temporal properties of the fading process; and c) increased link margins by simply collecting more of the transmitted energy at the receiver.

More recently, it has been realized that coordinated use of diversity can be achieved through the use of space–time codes. Rather than relying solely on array processing of uncoded transmissions, forward error correction codes which add redundancy in both the temporal and spatial domains are designed specifically for channels with multiple transmit and receive antennas. There are currently two main approaches to realizing the capacity potential of these channels: *coordinated* space–time codes and *layered* space–time codes.

Coordinated space-time block codes [11]–[13] and trellis codes [14]–[18] are designed for coordinated use in space and time. The data is encoded using multidimensional codes that span the transmit array. Trellis codes are typically decoded using the Viterbi algorithm. Such codes are efficient for small arrays, and can achieve within 3 dB of the 90% outage capacity rate calculated in [3]. A serious obstacle to extension to larger arrays, however, is the rapid growth of decoder complexity with array size and data rate: the number of states in a full-diversity space–time trellis code for t transmit antennas with rate R is $2^{(t-1)R}$. The other approach uses layered space–time codes [19]–[21], where the channel is decomposed into parallel single-input, single-output channels. The receiver successively decodes these layers by using antenna array techniques and linear or nonlinear cancellation methods. Both approaches show much promise, but the latter is more scalable. It also has the advantage that available technology such as standard error control codes can be more easily integrated.

Much of the space–time coding literature assumes the availability of good channel estimates, which are required for decoding. In the absence of channel knowledge, the capacity gains to be achieved depend upon the coherence time of the channel [22]–[25]. More recently, there has been considerable effort to design space–time codes that operate in the absence of channel state information [26]–[54]. Typically, these codes are coordinated space–time block codes whose symbols are unitary matrices. Such codes may also be used as part of differential space–time modulation schemes. Codes for use in differential schemes have been designed both using a group structure on the code and without such algebraic constraints.

Since space–time codes necessarily become rapidly very complex, there have been a number of papers proposing the concatenation of error control codes, mostly turbo codes, with a space–time code or space–time modulation [55]–[63]. This line of research is also motivated by the near-capacity-achieving performance of such concatenated codes on other channels. An overview in the area of concatenated space–time codes may be found in [64].

In [57], serial concatenation of standard space–time trellis codes with a bank of inner rate 1 recursive codes was considered. The outer code produces symbol vectors which are separately symbol interleaved. Each stream is sent through a separate inner recursive code with polynomial $1+D$. A quasi-static channel was used, which was assumed to be perfectly known at the receiver.

In [60], the authors consider space–time codes for two-transmit, two-receive antenna systems based on serial and parallel turbo codes. They develop a rank criteria for determination of the resulting diversity. They consider a parallel-concatenation-based approach in which a parallel turbo code is followed by puncturing and interleaving of each branch and transmission over separate antennas. For the serial approach [57], each output of a space–time trellis code is separately interleaved and fed into separate recursive encoders and transmitted over separate antennas. They consider both quasi-static fading and time-varying fading channel (0.001 and 0.01 normalized Doppler) and achieve 2 bits/s/Hz at 2.5 dB from outage capacity. Perfect channel knowledge was assumed.

In [55], [59], a form of turbo-coded space–time modulation is considered. The scheme uses serial concatenation of a standard turbo code with an interleaver, serial–parallel converter and a space–time modulator. The quasi-static fading channel is used and points 4 dB from capacity are achieved for a two-transmit, two-receive system. This is improved by about 2 dB with use of iterative demodulation. Separate channel estimation based on pilot sequences was also considered with a loss of about 1.5 dB.

In [65], the authors consider the parallel concatenation of recursive systematic space–time trellis codes to operate at 2 bits/s/Hz over a two-transmit two-receive antenna quasi-static fading channel. Antenna 1 transmits systematic symbols and antenna 2 transmits alternating parity symbols. Channel estimation was not considered and points 2.5 dB from outage capacity were achieved.

In [66], a rank criterion is developed and the authors consider the rate $1/3$ parallel turbo code from [67]. They design the interleaver to ensure full diversity and use binary phase-shift keying (BPSK) modulation to achieve 1 bit/s/Hz with either two or three transmit antennas and a single receive antenna. Channel estimation using pilot-symbol-aided modulation was considered resulting in a 2–3-dB loss in performance on fast fading channel with 0.01 and 0.001 normalized Doppler rates.

In [61], the authors design parallel concatenated space–time turbo codes based on constituent space–time trellis codes (in their recursive form). These codes are simulated over both fast fading and quasi-static fading channels with perfect channel knowledge.

In [63], a concatenation of standard convolutional codes with simple space–time block codes is considered, using bit interleaving. Perfect channel estimation is assumed and results are presented for quasi-static channels as well as fast fading channels with 0.01 and 0.05 normalized Doppler.

All of the approaches just described either assume perfectly known channel state information, or use simple channel estimation techniques in conjunction with training sequences or pilot symbols. Apart from rank criteria, no particular efforts are made to optimize the error control codes in these concatenated schemes.

In this correspondence, we show that the serial concatenation of a simple error control code and a differential space–time code of mod-

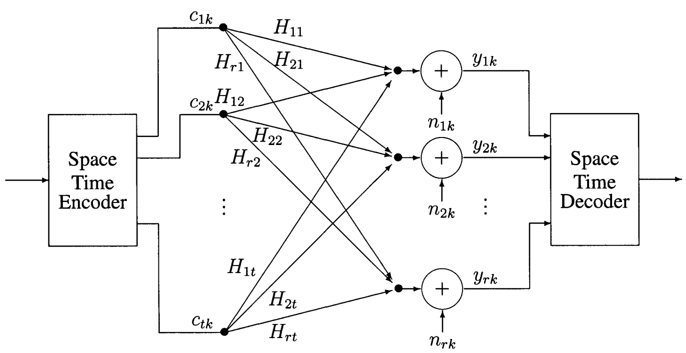

Fig. 1. A fading channel with multiple transmit and receive antennas.

erate size can practically achieve capacity on low signal-to-noise ratio (SNR) channels. This defines a decoder structure with manageable complexity and near-optimal performance. We pursue an approach motivated by turbo decoding methods. The main idea is to use a serially concatenated coding system in which the inner code is a short differential (recursive) space–time code. The outer code consists of a simple high-rate code, such as a convolutional code or even a parity-check code. Exploiting the turbo principle through iterative decoding, this arrangement can virtually achieve the Shannon limit of the multiple-antenna channel, even with an outer code so weak that it cannot correct single errors, and an inner code that is catastrophic. We further exploit the differential property of the inner code to provide *a priori* estimates of the fading channel, for use in an iterative channel estimation scheme. As the decoder iterates, both the quality of the data estimates and the channel estimates are improved. We find that this technique requires very few pilot symbols (since the transmitted data itself acts as soft, or uncertain training) and results in very little degradation compared to perfect channel knowledge.

Related independent work appearing subsequent to the initial publication of our methods [68], [69] can be found in [49], [50], [70]–[76]. Related independent prior work [77], using a similar serial concatenation of convolutional codes and differential space–time modulation, in conjunction with iterative decoding and per-survivor processing [78] was brought to our attention in the latter stages of the preparation of this correspondence.

The correspondence is organized as follows. In Section II, we introduce our mathematical model of the space–time channel. In Section III, we discuss the differential space–time modulation that will be used as an inner code in our concatenated system. Iterative decoding with perfect channel estimates is described in Section IV. In Section V, we use extrinsic information transfer charts to determine operation thresholds for these concatenated codes, which are compared with simulation in Section VI. In Section VII, we show how the differential space–time code can be used to provide initial *a priori* symbol estimates which are then used by an integrated channel estimator to refine channel estimates successively with the iterations of the decoder. The performance of this technique is shown using simulation in Section VIII. Section IX contains concluding remarks.

II. System Model

Transmission takes place over a channel with t transmit antennas and r receive antennas, as shown in Fig. 1. At time $k = 1, \ldots, n$, each transmit antenna $i = 1, \ldots, t$ selects a complex symbol c_{ik}, which is modulated onto a pulse waveform and transmitted over the channel. The vector $[c_{1k}, c_{2k}, \ldots, c_{tk}]$ is referred to as a *space–time symbol*. At each receive antenna $j = 1, \ldots, r$, the signal is passed through a filter matched to the pulse waveform and sampled synchronously. If the channel delay spread is negligible and fading conditions are approximately constant over n symbols, the samples taken by receive antenna j can be modeled as $y_{jk} = \sqrt{\rho/t}\sum_{i=1}^{t} H_{ji}c_{ik} + n_{jk}$, where H_{ji} is the complex fading path gain from the transmit antenna i to the receive antenna j, n_{jk} is a complex circularly symmetric Gaussian noise sample, and ρ is the SNR per receive antenna.

It is customary to collect the transmitted space–time symbols into a codeword matrix $\boldsymbol{C} \in \mathbb{C}^{t\times n}$ with elements c_{ik}, in which the rows correspond to different transmit antennas and the columns correspond to different symbol times. Considering a sequence of L codeword transmissions $\boldsymbol{C}_1, \boldsymbol{C}_2, \ldots, \boldsymbol{C}_L$, the channel can be written as

$$\boldsymbol{Y}_l = \sqrt{\frac{\rho}{t}}\boldsymbol{H}_l\boldsymbol{C}_l + \boldsymbol{N}_l \tag{1}$$

where $\boldsymbol{Y}_l \in \mathbb{C}^{r\times n}$ is the lth received matrix, $\boldsymbol{H}_l \in \mathbb{C}^{r\times t}$ is the matrix of fading path gains, and $\boldsymbol{N}_l \in \mathbb{C}^{r\times n}$ is a matrix of noise samples.

Independent Rayleigh fading may be modeled by selecting the elements of $\boldsymbol{H}_l$ as unit variance complex Gaussian random variables with independent and identically distributed (i.i.d.) real and imaginary parts. We distinguish between *fast* fading, in which the $\boldsymbol{H}_l$ evolve according to a process whose dominant frequency is much faster than $1/L$, but slower than $1/n$, and *quasi-static* fading, in which $\boldsymbol{H}_l$ is selected independently and then held constant for groups of L code matrices (corresponding to a single packet). In the fast fading case, it is mutual information averaged over the channel statistics that is used to calculate channel capacity, whereas for the quasi-static case, it is outage capacity that is of interest.

III. Differential Space–Time Turbo Codes

We shall focus on the differential space–time codes (DSTC) of Hughes [32]–[35], which can be demodulated without channel knowledge, at a loss of 3 dB in SNR with respect to decoding these codes with complete channel knowledge. These DSTCs are based on unitary matrices with a group structure, forming a space–time group code.[1] In such a code, each codeword takes the form $\boldsymbol{C} = \boldsymbol{D}\boldsymbol{Q}$, where $\boldsymbol{D}$ is a fixed $t \times n$ matrix and $\boldsymbol{Q} \in \mathcal{Q}$ belongs to a group (under matrix multiplication) of unitary matrices, $\boldsymbol{Q}\boldsymbol{Q}^* = \boldsymbol{I}$. The n columns of $\boldsymbol{C}$ are transmitted as n consecutive space–time symbols.

In particular, we shall illustrate the basic ideas for $t = n = 2$. For convenience, define the matrices

$$\boldsymbol{Q}^{(0)} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \qquad \boldsymbol{Q}^{(2)} = \begin{pmatrix} j & 0 \\ 0 & -j \end{pmatrix}$$
$$\boldsymbol{Q}^{(4)} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \qquad \boldsymbol{Q}^{(6)} = \begin{pmatrix} 0 & j \\ j & 0 \end{pmatrix}$$

and, correspondingly, $\boldsymbol{Q}^{(1)} = -\boldsymbol{Q}^{(0)}$, $\boldsymbol{Q}^{(3)} = -\boldsymbol{Q}^{(2)}$, $\boldsymbol{Q}^{(5)} = -\boldsymbol{Q}^{(4)}$, and $\boldsymbol{Q}^{(7)} = -\boldsymbol{Q}^{(6)}$. Then, the following defines a quaternary phase-shift keyed (QPSK) group code[2]

$$\mathcal{Q} = \left\{\boldsymbol{Q}^{(0)}, \boldsymbol{Q}^{(1)}, \ldots, \boldsymbol{Q}^{(7)}\right\} \tag{2}$$

$$\boldsymbol{D} = \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}. \tag{3}$$

DSTCs can be differentially encoded and decoded in a way very similar to phase-shift keying (PSK). At the start of transmission, the transmitter sends the code matrix $\boldsymbol{C}_0 = \boldsymbol{D}$. Thereafter, messages are differentially encoded: to send the information symbol $\boldsymbol{G}_l \in \mathcal{Q}$ during symbol time $l = 1, 2, \ldots,$ the transmitter sends

$$\boldsymbol{C}_l = \boldsymbol{C}_{l-1}\boldsymbol{G}_l. \tag{4}$$

[1]Group structure is not necessary for the design of a DSTC, and is not required for the decoding methods that we describe.

[2]As described in [33], $\mathcal{Q}$ is isomorphic to Hamilton's quaternion group.

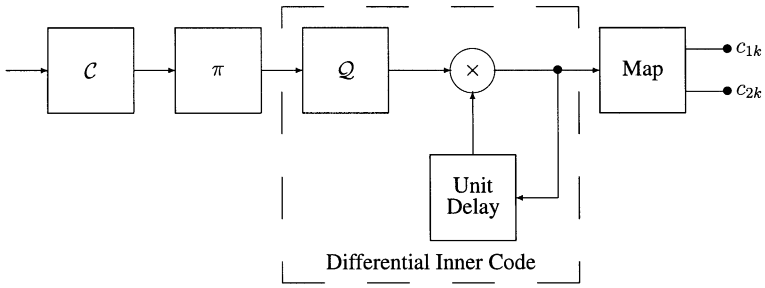

Fig. 2. Concatenation of an outer code with a differential space–time modulator.

The group property guarantees that $\boldsymbol{C}_l$ is a codeword if $\boldsymbol{C}_{l-1}$ is a codeword. Like differential PSK, there is a simple differential receiver [33] for $\boldsymbol{G}_l$ based on the two most recent blocks. This receiver computes

$$\hat{\boldsymbol{G}} = \max_{\boldsymbol{G}\in\mathcal{Q}} \Re \operatorname{tr} \boldsymbol{G}\boldsymbol{Y}_l^* \boldsymbol{Y}_{l-1}. \tag{5}$$

Performance results for the *quaternion code* (2) can be found in [33].

DSTCs provide an essential building block for space–time systems that can operate with or without channel estimates at the receiver. Thus far, most work on space–time coding has assumed that perfect channel estimates are available at the receiver [11], [19], [79]–[81]. In certain situations, however, the channel may change so rapidly as to make reliable estimation of the channel difficult, or may require too much overhead in terms of pilot symbols, or perhaps we want to avoid channel estimation in order to reduce the cost and complexity of the handset.

Fig. 2 shows the structure of the proposed serially concatenated code. For convenience, encoding is performed on a symbol-by-symbol level.

For each 4-ary input symbol, the rate $2/3$ outer code $\mathcal{C}$ outputs a single 8-ary symbol. For convenience, we shall consider these symbols to be members of $\mathbb{Z}_4 = \{0, 1, 2, 3\}$ and $\mathbb{Z}_8 = \{0, 1, \dots, 7\}$, respectively (although we do not invoke the ring axioms usually implied by this notation). We shall consider the standard maximal free distance convolutional codes [82, p. 495] with 4, 16, and 64 states (corresponding to $\nu = 1, 2,$ and 3) as well as the simple $(3, 2)$ parity-check code. The mapping from binary to $\mathbb{Z}_4$ and $\mathbb{Z}_8$ is not a topic of this correspondence and the natural mappings are used.

The stream of $\mathbb{Z}_8$ symbols are passed through a length L symbol interleaver π. These interleaved symbols are then input to the differential space–time encoder (the interleaved symbols $z \in \mathbb{Z}_8$ are mapped $\mathbb{Z}_8 \mapsto \mathcal{Q}$ according to $z \mapsto \boldsymbol{Q}^{(z)}$). In the case of the $(3, 2)$ parity code, bit interleaving must be used since the code has no dependencies between symbols. The mapper takes each 2×2 matrix output by the inner differential code and transmits it using two consecutive space–time symbols, as described earlier.

IV. Iterative Decoding With Channel Information

Since the differential space–time modulator is an infinite impulse response filter, it fits the requirements of an inner code for a serially concatenated coding system [83]. We now describe a symbol-wise turbo decoder [1] for the differential space–time turbo code, in the case that the receiver has perfect knowledge of the channel matrices $\boldsymbol{H}_l$. This decoder is of independent interest, since it forms the basis of the decoder for the case in which the channel matrices are unknown. We shall also see that using very simple constituent codes, we may achieve a performance very close to the fast-fading ergodic capacity for low SNRs.

The decoder is shown in Fig. 3. The inner *a posteriori* probability (APP) decoder operates on the (fully connected) trellis of the DSTC. This is followed by deinterleaving and APP decoding of the outer code. Extrinsic information on the space–time symbols is fed back to the DSTC decoder, which uses it as *a priori* information in a new decoding

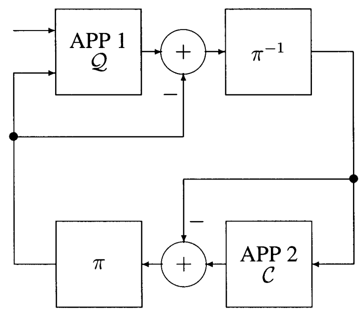

Fig. 3. Symbol-wise iterative decoder.

cycle. Such turbo decoding systems have been studied for the concatenation of standard convolutional codes with differential PSK, and very good error performance, rivaling that of turbo codes, have been reported [84], [85].

The inner decoder for the space–time differential code takes two inputs, the sequence of channel values $\{\boldsymbol{Y}_l\} = \{\boldsymbol{Y}_0, \boldsymbol{Y}_1, \dots\}$, and *a priori* information on the input symbols in the form of a sequence of probability vectors $\{\boldsymbol{p}_l^A\}$, where

$$\boldsymbol{p}_l^A = [\Pr(\boldsymbol{G}_l = \boldsymbol{Q}^{(0)}), \dots, \Pr(\boldsymbol{G}_l = \boldsymbol{Q}^{(M-1)})]$$

and $M = |\mathcal{Q}|$. It calculates the sequence of output APP vectors $\{\boldsymbol{p}_l\}$

$$\boldsymbol{p}_l = \left[\Pr\left(\boldsymbol{G}_l = \boldsymbol{Q}^{(0)} \middle| \{\boldsymbol{Y}_l\}\right), \dots, \Pr\left(\boldsymbol{G}_l = \boldsymbol{Q}^{(M-1)} \middle| \{\boldsymbol{Y}_l\}\right)\right] \tag{6}$$

using an adaptation of the standard implementation of the APP algorithm [86] (the differential constraint is represented in the usual way using the trellis upon which the decoder operates). Similarly, the APP algorithm can produce APPs on the coded symbols $\boldsymbol{C}_l$.

Given perfect channel knowledge, the algorithm uses the coherent branch metrics

$$\Pr(\boldsymbol{Y}_l \mid \boldsymbol{Q}) = \alpha \exp\left(-\|\boldsymbol{Y}_l - \boldsymbol{H}_l \boldsymbol{D}\boldsymbol{Q}\|_F^2\right) \tag{7}$$

where $\|\cdot\|_F$ is the Frobenius norm and α is a normalization factor.

The *a priori* information is removed from the output by element-wise division. The resulting "extrinsic" information $\{\boldsymbol{p}_l^E\}$ is given as

$$\boldsymbol{p}_l^E = \alpha \left[\boldsymbol{p}_l(1)/\boldsymbol{p}_l^A(1), \boldsymbol{p}_l(2)/\boldsymbol{p}_l^A(2), \dots, \boldsymbol{p}_l(M)/\boldsymbol{p}_l^A(M)\right]$$
$$\alpha^{-1} = \sum_{i=1}^{M} \boldsymbol{p}_l(i)/\boldsymbol{p}_l^A(i). \tag{8}$$

The extrinsic probabilities are deinterleaved in order to coincide with the symbols entering the outer APP decoder. This second decoder operates on the outer code $\mathcal{C}$. It uses $\{\boldsymbol{p}_l^E\}$ as *a priori* information, and calculates new APPs on the encoded symbols $\boldsymbol{G}_k$, according to the

APP rule. The new extrinsic probabilities are calculated using (8). This process is iterated until a certain stopping criterion is reached.

In case of very short outer codes such as the $(3, 2)$ parity-check code, bit interleaving needs to be used, and hence extrinsic probabilities for the information bits are generated by marginalizing the values obtained in (8). Likewise, the output bit probabilities of the outer code are combined into symbol probabilities.

V. Exit Analysis

In this section, we modify and apply the extrinsic information transfer analysis [87]–[90] to our serially concatenated system. This will give us a tool to predict the "turbo cliff" of our system and let us design and choose the outer code required for best performance.

Consider a single APP decoder. It operates on two inputs, the channel outputs $\boldsymbol{Y}_l$ (which may be absent if we consider the outer decoder), and *a priori* input information $\boldsymbol{p}_l^A$. The decoder generates extrinsic probabilities $\boldsymbol{p}_l^E$ according to (6) and (8).

Let the i.i.d. sequence of random variables $V_1, V_2, \ldots, V_L$ be the transmitted symbols. Under the assumption of long random interleavers, the corresponding sequence of random vectors $\boldsymbol{p}_l^A$ are i.i.d., as are the $\boldsymbol{p}_l^E$. Define the random variables V, A, and E according to $p(V, A) = p(V_l, \boldsymbol{p}_l^A)$ and $p(V, E) = p(V_l, \boldsymbol{p}_l^E)$ (note that these distributions are independent of l, due to the independence assumption). We can now define mutual informations $I(V; A)$ between the "true" symbols V and the input *a priori* probability vectors, and $I(V; E)$ between V and the output extrinsic probability vectors. Note that this amounts to treating the prior and extrinsic probability vectors themselves as the random vectors of interest (as opposed to considering the random variables taking these vectors as their distributions). This makes sense, since it is from these vectors that the final decisions will be made.

The plot of $I(V, A)$ versus $I(V, E)$ is known as extrinsic information transfer (EXIT) chart. By plotting the EXIT charts for the inner and outer code on the same axes, the convergence properties of the concatenated system may be predicted, as described in [87]–[90].

Due to the nonlinear operation of the APP decoder, the multidimensional distributions $p(V, A)$ and $p(V, E)$ are difficult to obtain analytically. We, therefore, estimate these measures using Monte Carlo simulation of the individual codes of interest (the DSTC and the convolutional code). Further details regarding symbol-wise EXIT analysis may be found in [91].

Fig. 4 shows the EXIT chart for our system using the DSTC as the inner code ($I_A = I(V, A)$ on the horizontal axis, $I_E = I(V, E)$ on the vertical axis) and 4-, 16-, and 64-state maximal free distance rate $2/3$ convolutional codes [82, p. 495] as the outer code (thick lines) (I_A on the vertical axis, I_E on the horizontal axis). The DSTC curves (thin lines) are for various SNRs, ranging from -1.5 dB to -1.0 dB in steps of 0.1 dB.

From the figure, we expect the turbo threshold to occur at about -1.2 dB for the 64-state outer code, -1.3 dB for the 16-state code, and -1 dB for the 4-state code. We also expect the 16- and 64-state outer codes to result in faster convergence, since the path between the inner and outer curves are more open for these codes.

The $(3, 2)$ parity-check code (dashed line) has an even lower turbo cliff at about -1.4 dB. Its curve is very well matched to the shallow curve of the inner decoder. Furthermore, the APP decoder for this parity-check code is extremely simple, and can be implemented by a simple lookup table. From the parity-check constraint it is quite straightforward to calculate the output extrinsic bit log-likelihood ratio as

$$\lambda^E(b_1) = \lambda^A(b_2) + \log\left(\frac{1+\exp(\lambda^A(b_3)-\lambda^A(b_2))}{1+\exp(\lambda^A(b_3)+\lambda^A(b_2))}\right) \quad (9)$$

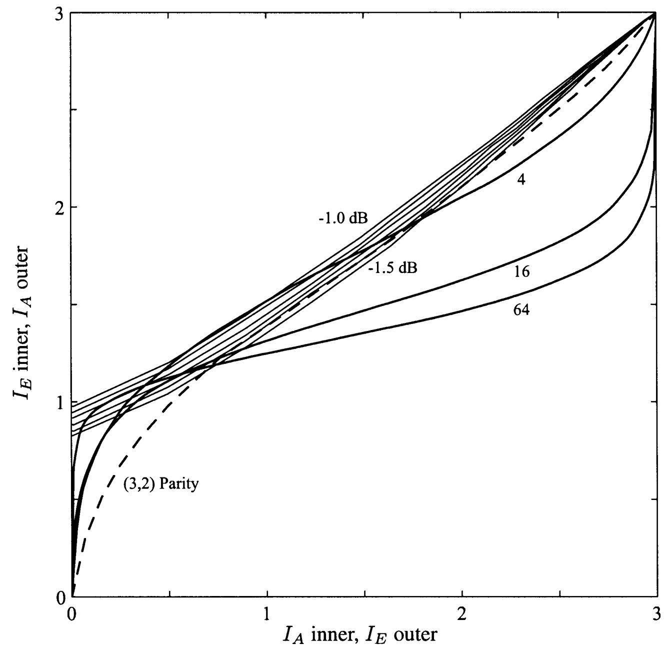

Fig. 4. Extrinsic information transfer chart for serial concatenation of 4-, 16-, 64-state convolutional codes and the $(3, 2)$ parity-check code with the DSTC.

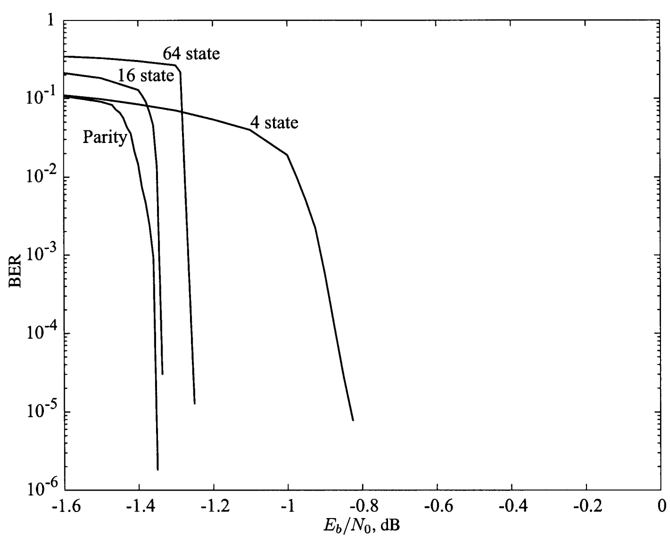

Fig. 5. Performance of the serially concatenated system with several outer codes for a fast fading channel.

where b_1, b_2, b_3 are the bits that make up a single parity-check codeword. It is clear however that many iterations may be required for convergence.

VI. Simulation Results

Fig. 5 compares the performance of the differential space–time turbo-coded system using the various outer codes previously discussed. The channel used is assumed to be a $t = r = 2$ uncorrelated fast Rayleigh-fading channel, i.e., each channel matrix $\boldsymbol{H}_l$ is independently drawn from a 2×2 matrix of independent complex Gaussian distributions, and the decoder is furnished with ideal channel side information. The interleaver length was 100 000 symbols, and 100 decoding iterations were performed (smaller numbers of iterations result in less steep curves).

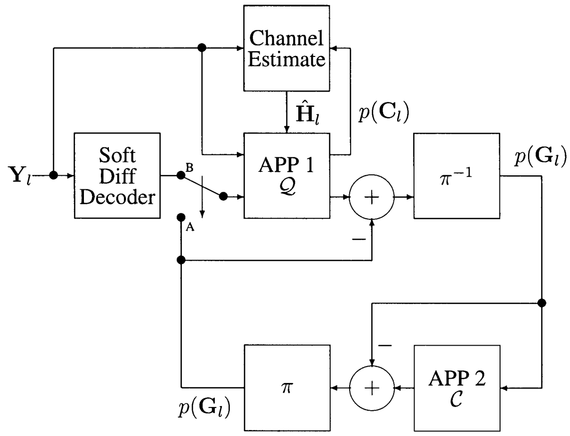

Fig. 6. Symbol-wise iterative decoder with integrated channel estimation.

The 4-, 16-, and 64-state outer convolutional codes result in thresholds of approximately -1, -1.35, and -1.25 dB, respectively, which agree with the predictions of the EXIT analysis. The threshold for the $(3, 2)$ parity code is similar to that of the 16-state convolutional code, at about -1.35 dB.

At this code rate (1 bit per channel use), and for $t = r = 2$ capacity is at -3.1 dB [2]. Thus, the concatenated coding system can achieve bit error rates (BERs) of 10^{-6} at about 1.75 dB from capacity (for the parity outer code). However, this code also requires the most iterations as predicted by the EXIT analysis. By way of comparison, the 16-state convolutional code achieves a BER of 10^{-6} at $E_b/N_0 = -1.3$ dB after 25 iterations, whereas the parity code requires 45 iterations to achieve the same point.

We may also compare these results to those for (nonconcatenated) space–time block codes. As reported in [64, Fig. 10], the 1-bit/s/Hz BPSK Alamouti code [11] with $t = r = 2$ achieves a BER of 10^{-3} at $E_b/N_0 = 7$ dB. The corresponding 1-bit/s/Hz QPSK code with $t = 4$, $r = 2$ from [12] achieves a BER of 10^{-3} at about $E_b/N_0 = 5$ dB.

In [64, Fig. 26], results are given for turbo space–time codes formed by the concatenation of standard convolutional codes with the QPSK Almouti code [11] resulting in 1/bit/s/Hz. Using a constraint length 5 outer code, a BER of 10^{-3} is achieved at $E_b/N_0 = 4.5$ dB. Increasing the constraint length of the outer code to 9 improves this result to $E_b/N_0 = 3.25$ dB.

VII. Iterative Decoding Without Channel Information

We now describe an iterative decoder that makes use of the differential property of the inner code to begin. This iterative decoder is shown in Fig. 6.

In the case that the channel is not known, the differential property of the inner code may be used to provide initial (nonuniform) priors to the inner APP decoder on the space–time symbols $\boldsymbol{G}_l$ via

$$P(\boldsymbol{G}_l|\boldsymbol{Y}_l) \propto \exp(\Re \operatorname{tr} \boldsymbol{G}\boldsymbol{Y}_l^*\boldsymbol{Y}_{l-1}). \tag{10}$$

This is the function of the soft differential decoder in Fig. 6.

With reference to Fig. 6, the switch is held at position *B* for the first iteration and the inner APP decoder operates on the symbol priors (10), ignoring the channel symbols $\boldsymbol{Y}_l$, since these contain no useful information without channel knowledge. The decoder APP 1 generates channel symbol posteriors $p(\boldsymbol{C}_l)$ which are used to form channel estimates to be used in subsequent iterations by a channel estimator discussed below.

It is easy to see that the symbol posteriors $p(\boldsymbol{G}_l)$ at the output of decoder APP 1 are identical to its input (10), i.e., without channel knowledge APP 1 cannot improve the symbol probabilities. These are deinterleaved and fed into the outer decoder APP 2, which generates new symbol priors $p(\boldsymbol{G}_l)$. At this point, further reuse of the differential decoder can be shown to be of no use, and the system switches to "coherent" operation in the next iteration. With the switch in position *A*, priors $p(\boldsymbol{G}_l)$ are used from the previous iteration, together with the channel estimates $\hat{\boldsymbol{H}}_l$ using the coherent metric (7) (treating the channel estimates as if they were exact). At each iteration, improved values $p(\boldsymbol{C}_l)$ are also generated by the inner decoder APP 1 which, in turn, generates improved channel estimates $\hat{\boldsymbol{H}}_l$.

The transmission channel (1) is linear, and therefore well-documented optimal linear estimators exist. In particular, the linear minimum mean-square error (LMMSE) estimator can be formulated as $\hat{\boldsymbol{H}} = \boldsymbol{R}_{HY}\boldsymbol{R}_{YY}^{-1}\boldsymbol{Y}$, where $\hat{\boldsymbol{H}} = [\hat{\boldsymbol{H}}_1, \ldots, \hat{\boldsymbol{H}}_L]$, and $\boldsymbol{Y} = [\boldsymbol{Y}_1, \ldots, \boldsymbol{Y}_L]$. The covariance matrices are given by $\boldsymbol{R}_{HY} = E[\boldsymbol{H}\boldsymbol{Y}^*]$ and $\boldsymbol{R}_{YY} = E[\boldsymbol{Y}\boldsymbol{Y}^*]$. These covariance matrices need to be estimated using not only the statistics of the channel, which are Gaussian, but also the statistics of the transmitted symbols $\boldsymbol{C}_l$, whose individual symbol probabilities are furnished by decoder APP 1 at each iteration.

Even a cursory complexity study of this LMMSE filtering approach reveals that it would dominate the complexity of the entire receiver, especially given the very simple soft component decoders for the small error control codes used. Motivated by [92]–[94], we pursue a simpler estimator, in particular, we reason that from the first moment equation $E[\boldsymbol{Y}_l] = \sqrt{\rho/t}\ \boldsymbol{H}_l E[\boldsymbol{C}_l]$, an instantaneous symbol-wise estimate of the channel is found as

$$\sqrt{\frac{\rho}{t}}\,\tilde{\boldsymbol{H}}_l = \boldsymbol{Y}_l\tilde{\boldsymbol{C}}_l^* \tag{11}$$

where $\tilde{\boldsymbol{C}}_l$ is the empirical mean according to $p(\boldsymbol{C}_l)$ generated by the inner decoder APP 1.

The symbol-wise estimates $\tilde{\boldsymbol{H}}_l$ are now treated as if they were in fact observations of $\boldsymbol{H}_l$ in i.i.d. additive Gaussian noise. The final channel estimates are obtained by LMMSE filtering the elements of $\tilde{\boldsymbol{H}}_l$ according to the known or estimated covariance of the fading process. In the case of quasi-static fading this amounts to averaging.

The *a posteriori* channel symbol probabilities used to generate the average symbol $\tilde{\boldsymbol{C}}_l$ used in (11) are provided by the inner decoder APP 1. However, due to the catastrophic nature of the quaternion code (and all differential codes), even completely reliable values of the input symbols $\boldsymbol{G}_l$ cannot give any information on the channel symbols $\boldsymbol{C}_l$. But if the quaternion code is started in a known state, nonuniform probabilities $p(\boldsymbol{C}_l)$ of the channel symbols are found by decoder APP 1. We show now that these probabilities "decay" to the uniform distribution with a decay rate that is a function of only the channel SNR and the differential code used, but does not directly depend on the channel fluctuation speed.

The forward recursion of the inner APP decoder can be written in matrix notation as $\boldsymbol{\alpha}_{i+1} = \boldsymbol{P}_i\boldsymbol{\alpha}_i$, where $\boldsymbol{\alpha}_i$ is a state-size vector of s forward recursion values, and $\boldsymbol{P}_i$ is the transition probability matrix between the states, whose (k, l)th entry is the transition probability between state k and state l given by $\Pr(Y|\boldsymbol{C})p(G)$, where $\boldsymbol{G}$ is the input symbol and $\boldsymbol{C}$ the channel symbol on the branch from state k to l. After l steps we have

$$\boldsymbol{\alpha}_{i+l} = \left(\prod_{j=i}^{i+l} \boldsymbol{P}_i\right)\boldsymbol{\alpha}_i. \tag{12}$$

The product on the right-hand side is dominated by the largest eigenvalues of the $\boldsymbol{P}_i$, and since $\boldsymbol{P}_i$ is a doubly stochastic matrix for the quaternion code, it has, independent of i, the largest eigenvalue $\lambda_1 = 1$

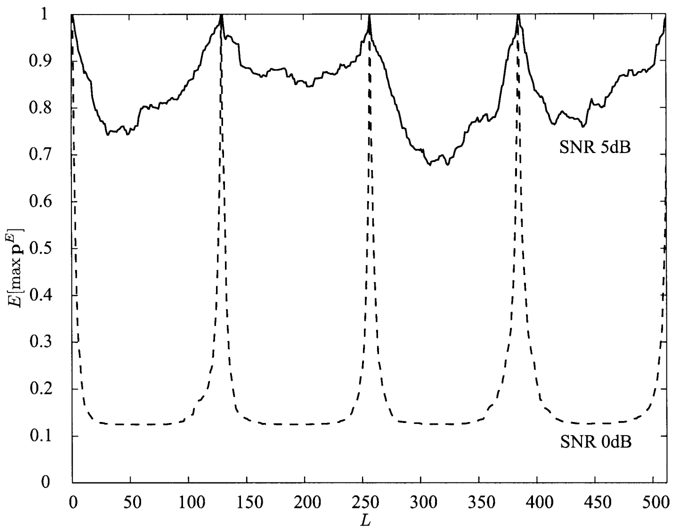

Fig. 7. Training memory effect.

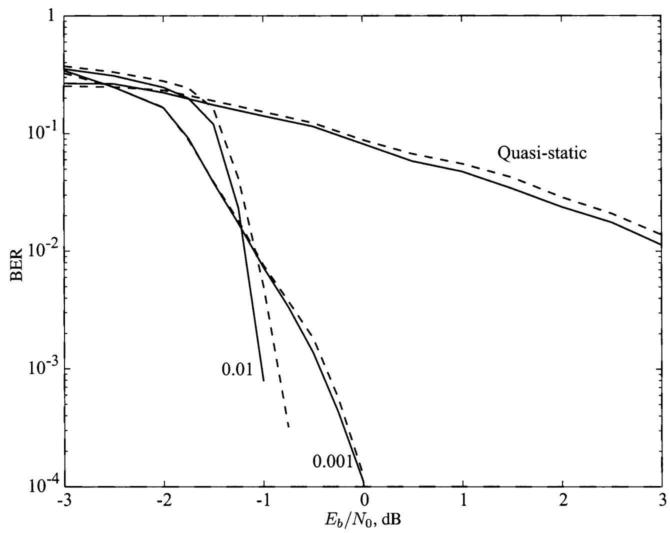

Fig. 8. Performance of the noncoherent iterative decoder.

with the all-ones eigenvector $\boldsymbol{e}$. Writing out the spectral decomposition of $\boldsymbol{P}_i$ as

$$\boldsymbol{P}_i = \boldsymbol{e}\boldsymbol{e}'/s + \sum_{r=2}^{s} \lambda_{i,r} \boldsymbol{u}_{i,r} \boldsymbol{v}_{i,r}^*$$

where $\boldsymbol{u}_r$ and $\boldsymbol{v}_r$ are the left and right eigenvectors to λ_r, we see that, as $l \to \infty$, $\lambda_1 = 1$ will dominate in the matrix product (12), and $\boldsymbol{\alpha}_{i+l} \to [1/s, \ldots, 1/s]'$, regardless of the initial value of $\boldsymbol{\alpha}_i$. The speed with which this uniform distribution is approached is essentially determined by $\lambda_{i,2}$, the second largest eigenvalues of $\boldsymbol{P}_i$, and their values depend on the channel SNR. This persistence of the initial state information in the form of nonuniform channel probabilities is called *training memory*.

Since the effects of an initial known state die out, periodical insertion of known input symbols can be used to force the encoder into a known state and thus refresh the training memory. To preserve rate, the channel symbols corresponding to these reset symbols are punctured. The required rate of these refresh symbols depends only on the channel SNR. As long as the training memory is never allowed to decay too much, fading processes with large bandwidth may be estimated using the memory of the code. In the case of quasi-static channels with high enough SNR, reset symbols may not be required at all, since the channel estimate obtained from the start of the packet may be used for the entire packet.

VIII. Simulation Results Without Channel Information

We now present the results of computer simulation for the case when the channel information $\boldsymbol{H}$ is unknown to the receiver. First, we demonstrate the training memory effect for the $t = r = 2$ quasi-static channel. Fig. 7 shows how the APPs at the output of APP 1 decay as a function of the channel SNR. The horizontal axis is the symbol time and the vertical axis shows the sample average of the maximum APP. Refresh symbols were inserted every 128 symbols. From the figure we see that at 0 dB (dashed line), the distribution decays very quickly to uniform (within a few symbols). In contrast, at 5 dB, the training memory keeps the probability of the correct symbol from decaying below about 0.7.

Fig. 8 shows simulation results for the decoder of Fig. 6. The 16-state outer convolutional code was used, with blocks of 5000 symbols and 50 decoder iterations (using the parity code as the outer code results in similar performance). Three $t = r = 2$ fading channels have been simulated, a quasi-static channel, and two low-pass channels. For the low-pass channels, the complex gains between each antenna pair are independent low-pass filtered complex Gaussian processes. Normalized (to the symbol rate) fading bandwidths of $f_c = 0.01$ and $f_c = 0.001$ were considered. For this channel model, the LMMSE channel estimator described in the previous section amounts to individual filtering of the independent processes, according to the fading bandwidth and the noise variance (which are assumed known). For the quasi-static and $f_c = 0.001$ channels, the training memory was refreshed every 128 symbols. For the faster $f_c = 0.01$ channel, a refresh period of 32 symbols was used. Bear in mind that the operating point of this system is at a symbol SNR of about 0 dB, and from Fig. 7 we saw that for 0 dB, the extrinsic probability decays to uniform very quickly. The slower channels could make do with such a slow refresh rate because the few "good" symbols obtained from each refresh are enough to estimate the channel at that point, and the channel does not change too much prior to the next refresh. Profiting from the channel memory, fewer refresh symbols are needed than required by training memory alone. For the faster channel, we need to make sure that the training memory does not decay much at all between refresh symbols.

For each case, the BER performance of the iterative receiver with perfect channel knowledge is shown with a solid line, and the performance of the iterative receiver without channel knowledge is shown with a dashed line. In all cases, we see that the latter obtains a performance very close to that of the coherent receiver, within approximately 0.1 dB. This can be compared to the results of [66], where they reported a loss of 2–3 dB due to channel estimation using conventional techniques (for $r = 1$ and $t = 2, 3$ systems at 1 bit/s/Hz).

IX. Conclusion

We have shown that the serial concatenation of simple convolutional or block codes with differential space–time modulation can provide outstanding performance over the multiple-antenna channel, using symbol-based turbo decoding. We have presented an EXIT chart analysis which guides the choice of the optimal component codes for a given situation and demonstrated the tightness of this analysis with example simulations. The concatenation of standard convolutional and block codes with the inner differential code yields space–time turbo codes that are interesting in their own right. These codes provide a gain of over 8 dB compared to the Alamouti block space–time code [11] and over 6 dB compared to the Tarokh *et al.* code [12] (which uses twice as many transmit antennas).

We also have shown how the differential inner code can be exploited to provide a simple integrated channel estimator with soft information to generate channel estimates which are reliable enough to allow iterative decoding and channel estimation to start up. Performance results on both block fading and fast fading channels demonstrate that such a receiver with this integrated channel estimation is effective and can achieve the performance of a system with ideal channel information to within a fraction of a decibel.

References

[1] C. Berrou and A. Glavieux, "Near optimum error correcting coding and decoding: Turbo-codes," *IEEE Trans. Commun.*, vol. 44, pp. 1261–1271, Oct. 1996.

[2] I. E. Telatar, "Capacity of multi-antenna Gaussian channels," *Europ. Trans. Telecomm.*, vol. 10, no. 6, pp. 585–595, Nov.–Dec. 1999. Origninally published as Tech. Memo., Bell Labs., Lucent Technologies, Oct. 1995.

[3] G. J. Foschini and M. J. Gans, "On limits of wireless communications in a fading environment when using multiple antennas," *Wireless Personal Commun.*, vol. 6, no. 3, pp. 311–335, Mar. 1998.

[4] J. H. Winters, J. Salz, and G. D. Gitlin, "The capacity of wireless communication systems can be substantially increased by the use of antenna diversity," in *Proc. Conf. Information Science and Systems*, Princeton, NJ, 1992.

[5] J. H. Winters, J. Salz, and R. Gitlin, "The impact of antenna diversity on the capacity of wireless communication systems," *IEEE Trans. Commun.*, vol. 42, pp. 1740–1751, Feb./Mar./Apr. 1994.

[6] A. Narula, M. D. Trott, and G. W. Wornell, "Performance limits of coded diversity methods for transmitter antenna arrays," *IEEE Trans. Inform. Theory*, vol. 45, pp. 2418–2433, Nov. 1999.

[7] S. Dennett, "The cdma2000 ITU-R RTT candidate submission (0.18)," TIA, Tech. Rep., July 1998.

[8] W. C. Jakes, *Microwave Mobile Communications*. New York: Wiley, 1974.

[9] D. H. Johnson and D. E. Dudgeon, *Array Signal Processing: Concepts and Techniques*. Englewood Cliffs, NJ: Prentice-Hall, 1993.

[10] S. Verdú, *Multiuser Detection*. Cambridge, U.K.: Cambridge Univ. Press, 1998.

[11] S. M. Alamouti, "A simple transmit diversity technique for wireless communications," *IEEE J. Select. Areas Commun.*, vol. 16, pp. 1451–1458, Oct. 1998.

[12] V. Tarokh, H. Jafarkhani, and A. R. Calderbank, "Space-time block codes from orthogonal designs," *IEEE Trans. Inform. Theory*, vol. 45, pp. 1456–1467, July 1999.

[13] ——, "Space-time block coding for wireless communications: Performance results," *IEEE J. Select. Areas Commun.*, vol. 17, pp. 451–460, Mar. 1999.

[14] 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.

[15] A. F. Naguib, V. Tarokh, N. Seshadri, and A. 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.

[16] D. Agrawal, V. Tarokh, A. Naguib, and N. Seshadri, "Space-time coded OFDM for high data-rate wireless communication over wideband channels," in *Proc. 48th IEEE Vehicular Technology Conf.*, vol. 3, Ottawa, ON, Canada, May 1998, pp. 2232–2236.

[17] V. Tarokh, A. Naguib, N. Seshadri, and A. R. Calderbank, "Combined array processing and space-time coding," *IEEE Trans. Inform. Theory*, vol. 45, pp. 1121–1128, May 1999.

[18] S. Baro, G. Bauch, and A. Hansmann, "Improved codes for space-time trellis-coded modulation," *IEEE Commun. Lett.*, vol. 4, pp. 20–22, Jan. 2000.

[19] 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, no. 2, pp. 41–159, Aug. 1996.

[20] P. W. Wonianski, G. J. Foschini, G. D. Golden, and R. A. Valenzuela, "V-BLAST: An architecture for realizing very high data rates over rich scattering wireless channels," in *Proc. ISSSE*, 1998, pp. 295–300.

[21] H. El Gamal and A. R. Hammons, "A new approach to layered space-time coding and signal processing," *IEEE Trans. Inform. Theory*, pp. 2321–2334, Sept. 2001.

[22] T. L. Marzetta and B. M. Hochwald, "Capacity of a mobile multiple-antenna communication link in Rayleigh flat fading," *IEEE Trans. Inform. Theory*, vol. 45, pp. 139–157, Jan. 1999.

[23] B. Hassibi and B. M. Hochwald, "How much training is needed in multiple-antenna wireless links?," *IEEE Trans. Inform. Theory*, pp. 951–963, Apr. 2003.

[24] L. Zheng and D. Tse, "Sphere packing in the Grassman manifold: A geometric approach to the noncoherent multi-antenna channel," in *IEEE Int. Symp. Information Theory*, Sorrento, Italy, June 2000.

[25] L. Zheng and D. N. C. Tse, "Communication on the Grassmann manifold: A geometric approach to the noncoherent multiple-antenna channel," *IEEE Trans. Inform. Theory*, vol. 48, pp. 359–383, Feb. 2002.

[26] B. Hochwald and W. Sweldens, "Differential unitary space-time modulation," *IEEE Trans. Commun.*, vol. 48, pp. 2041–52, Dec. 2000.

[27] T. L. Marzetta, B. Hochwald, and B. Hassibi, "Space-time autocoding: Arbitrarily reliable communication in a single fading interval," in *Proc. IEEE Int. Symp. Information Theory*, Sorrento, Italy, June 2000, p. 313.

[28] D. Agrawal, T. J. Richardson, and R. Urbanke, "Multiple-antenna signal constellations for fading channels," in *Proc. IEEE Int. Symp. Information Theory*, Sorrento, Italy, June 2000, p. 365.

[29] ——, "Multiple-antenna signal constellations for fading channels," *IEEE Trans. Inform. Theory*, vol. 47, pp. 2618–2626, Sept. 2001.

[30] B. M. Hochwald and T. L. Marzetta, "Unitary space-time modulation for multiple-antenna communications in Rayleigh flat fading," *IEEE Trans. Inform. Theory*, vol. 46, pp. 543–564, Mar. 2000.

[31] B. M. Hochwald, T. L. Marzetta, T. J. Richardson, W. Sweldens, and R. Urbanke, "Systematic design of unitary space-time constellations," *IEEE Trans. Inform. Theory*, vol. 46, pp. 1962–1973, Sept. 2000.

[32] B. Hughes, "Extensions to the theory of differential space-time modulation," in *Proc. IEEE Int. Symp. Information Theory*, Sorrento, Italy, June 2000, p. 285.

[33] B. L. Hughes, "Differential space-time modulation," *IEEE Trans. Inform. Theory*, vol. 46, pp. 2567–2578, Nov. 2000.

[34] ——, "Optimal space-time constellations from groups," *IEEE Trans. Inform. Theory*, vol. 46, pp. 401–410, Feb. 2000.

[35] ——, "Space-time coding without channel estimation," in *Proc. Conf. Information Science and Systems*, Johns Hopkins Univ., Baltimore, MD, Mar. 22–24, 1999.

[36] ——, "Differential space-time modulation," in *Proc. IEEE Wireless Commun. Networking Conf.*, New Orleans, LA, Sept. 22–29, 1999, pp. 145–9.

[37] ——, "Further results on differential space-time modulation," in *Proc. IEEE Sensor Array and Multichannel Signal Processing Workshop*, Cambridge, MA, Mar. 16–17, 2000.

[38] X.-B. Liang and X.-G. Xia, "Some unitary signal constellations for differential space-time modulation," in *Proc. 34th Asilomar Conf. Signals, Systems and Computers*, vol. 1, 2000, pp. 716–720.

[39] B. Hassibi, B. Hochwald, A. Shokrollahi, and W. Sweldens, "Codes for differential signaling with many antennas," in *Proc. Wireless Communications and Networking Conf. (WCNC)*, vol. 1, Chicago, IL, Sept. 2000, pp. 23–24.

[40] B. M. Hochwald and T. L. Marzetta, "Space-time autocoding," *IEEE Trans. Inform. Theory*, vol. 47, pp. 2761–2781, Nov. 2001.

[41] X.-B. Liang and X.-G. Xia, "A class of two by two unitary signal constellations for differential space-time modulation," in *Proc. IEEE Int. Symp. Information Theory*, Washington, D.C., June 24–29, 2001, p. 81.

[42] A. Shokrollahi, "Group characters and unitary space-time codes," in *Proc. IEEE Int. Symp. Information Theory*, Washington, D.C., June 24–29, 2001, p. 107.

[43] ——, "Design of unitary space-time codes from representations of $SU(2)$," in *Proc. IEEE Int. Symp. Information Theory*, Washington, D.C., June 24–29, 2001, p. 241.

[44] J. Liu, J. Li, H. Li, and E. G. Larsson, "Differential space-code modulation for interference suppression," *IEEE Trans. Signal Processing*, vol. 49, pp. 1786–1795, Aug. 2001.

[45] J. Liu, E. G. Larsson, J. Li, and H. Li, "High-rate differential space-code modulation for interference suppression," in *IEEE 3rd Workshop on Signal Processing Advances Wireless Communications*, 2001, pp. 283–286.

[46] C. B. Peel and A. L. Swindlehurst, "Performance of unitary space-time modulation in a continuously changing channel," in *Proc. IEEE Int. Conf. Communications (ICC 2001)*, vol. 9, 2001, pp. 2805–2808.

[47] M. Tao and R. S. Cheng, "Differential space-time block codes," in *IEEE Global Telecommunications Conf.*, San Antonio, TX, 2001, pp. 1098–1102.

[48] G. Ganesan and P. Stoica, "Differential modulation using space-time block codes," *IEEE Signal Processing Lett.*, vol. 9, pp. 57–60, Feb. 2002.

[49] L. H.-J. Lampe and R. Schober, "Bit-interleaved coded differential space-time modulation," in *Proc. IEEE Int. Conf. on Communications (ICC 2002)*, vol. 3, 2002, pp. 1434–1438.

[50] ——, "Bit-interleaved coded differential space-time modulation," *IEEE Trans. Commun.*, vol. 50, pp. 1429–1439, Sept. 2002.

[51] S. N. Diggavi, N. Al-Dhahir, A. Stamoulis, and A. R. Calderbank, "Differential space-time coding for frequency-selective channels," *IEEE Commun. Lett.*, vol. 6, pp. 253–255, June 2002.

[52] X.-B. Liang and X.-G. Xia, "Unitary signal constellations for differential space-time modulation with two transmit antennas: Parametric codes, optimal designs, and bounds," *IEEE Trans. Inform. Theory*, vol. 48, pp. 2291–2322, Aug. 2002.

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

[54] T. L. Marzetta, B. Hassibi, and B. M. Hochwald, "Structured unitary space-time autocoding constellations," *IEEE Trans. Inform. Theory*, vol. 48, pp. 942–950, Apr. 2002.

[55] A. Stefanov and T. M. Duman, "Turbo coded modulation for wireless communications with antenna diversity," in *Proc. IEEE Vehicular Technology Conf.*, vol. 3, Sept. 1999, pp. 1565–1569.

[56] K. R. Narayanan, "Turbo decoding of concatenated space-time codes," in *Proc. 37th Allerton Conf. Communications, Control, and Computing*, Monticello, IL, Sept. 1999.

[57] X. Lin and R. S. Blum, "Improved space-time codes using serial concatenation," *IEEE Commun. Lett.*, vol. 4, pp. 221–223, July 2000.

[58] I. Bahceci and T. M. Duman, "Combined turbo coding and unitary space-time modulation," in *Proc. IEEE Int. Symp. Information Theory*, Washington, DC, June 2001, p. 106.

[59] A. Stefanov and T. Duman, "Turbo-coded modulation for systems with transmit and receive antenna diversity over block fading channels: System model, decoding approaches, and practical considerations," *IEEE J. Select. Areas Commun.*, vol. 19, pp. 958–968, May 2001.

[60] Y. Liu, M. P. Fitz, and O. Y. Takeshita, "Full rate space-time turbo codes," *IEEE J. Select. Areas Commun.*, vol. 19, pp. 969–980, May 2001.

[61] J. Yuan, B. Vucetic, Z. Chen, and W. Firmanto, "Design of space-time turbo tcm on fading channels," in *Proc. 2001 IEEE Information Theory Workshop*, Cairns, Australia, 2001, pp. 123–125.

[62] W. Firmanto, Z. Chen, B. Vucetic, and J. Yuan, "Design of space-time turbo trellis coded modulation for fading channels," in *Proc. Global Telecommunications Conf. (GLOBECOM '01)*, vol. 2, 2001, pp. 1093–1097.

[63] Z. Hong and B. L. Hughes, "Robust space-time codes for time selective fading," in *Proc. IEEE Information Theory Workshop*, Cairns, Australia, Sept. 2001, pp. 112–114.

[64] T. H. Liew and L. Hanzo, "Space-time codes and concatenated channel codes for wireless communications," *Proc. IEEE*, vol. 90, pp. 187–219, Feb. 2002.

[65] D. Cui and A. M. Haimovich, "Performance of parallel concatenated space-time codes," *IEEE Commun. Lett.*, vol. 5, pp. 236–238, June 2001.

[66] H.-J. Su and E. Geraniotis, "Space-time turbo codes with full antenna diversity," *IEEE Trans. Commun.*, vol. 49, pp. 47–57, Jan. 2001.

[67] A. R. Hammons and H. El Gamal, "On the theory of space-time codes for PSK modulation," *IEEE Trans. Inform. Theory*, vol. 46, pp. 524–542, Mar. 2000.

[68] A. Grant and C. Schlegel, "Differential turbo space-time coding," in *Proc. IEEE Information Theory Workshop*, Cairns, Australia, Sept. 2–7, 2001, pp. 120–122.

[69] C. Schlegel and A. Grant, "Concatenated space-time coding," in *Proc. Symp. Personal, Indoor and Mobile Radio Communications (PIMRC 2001)*, San Diego, CA, 2001, pp. 139–143.

[70] R. Schober and L. H.-J. Lampe, "Noncoherent receivers for differential space-time modulation," in *Proc. IEEE Global Telecommunications Conf.*, vol. 2, 2001, pp. 1127–1131.

[71] A. Steiner, M. Peleg, and S. Shamai, "Turbo coded space-time unitary matrix differential modulation," in *Proc. IEEE VTS 53rd Vehicular Technology Conf. (VTC 2001) Spring.*, vol. 2, 2001, pp. 1352–1356.

[72] C. Ling and X. Wu, "Linear prediction receiver for differential space-time modulation over time-correlated rayleigh fading channels," in *Proc. IEEE Int. Conf. Communications*, vol. 2, 2002, pp. 788–791.

[73] R. Schober and L. H.-J. Lampe, "Noncoherent receivers for differential space-time modulation," *IEEE Trans. Commun.*, vol. 50, pp. 768–777, May 2002.

[74] A. Steiner, M. Peleg, and S. Shamai (Shitz), "Iterative decoding of space-time differentially coded unitary matrix modulation," *IEEE Trans. Signal Processing*, vol. 50, pp. 2385–2395, Oct. 2002.

[75] Y. Yao and M. Howlader, "Serial and parallel concatenated differential space-time coding," in *Proc. MILCOM*, vol. 1, 2002, pp. 721–725.

[76] P. Tarasak and V. K. Bhargava, "Reduced complexity multiple symbol differential detection of space-time block code," in *Proc. Wireless Communications and Networking Conf. (WCNC2002).*, Mar. 17–21, 2002, pp. 505–509.

[77] A. V. Nguyen and M. A. Ingram, "Iterative demodulation and decoding of differential space-time block codes," in *Proc. IEEE Vehicular Technology Conf.*, Boston, MA, 2000, pp. 2394–2400.

[78] R. Raheli, A. Polydoros, and C.-K. Tzou, "Per-survival processing: A general approach to MLSE in uncertain environments," *IEEE Trans. Commun.*, vol. 43, pp. 354–364, Feb./Mar./Apr. 1995.

[79] N. Seshadri and J. H. Winters, "Two signaling schemes for improving the error performance of frequency-division-duplex (FDD) transmission systems using transmitter antenna diversity," *Int. J. Wireless Inform. Networks*, vol. 1, no. 1, 1994.

[80] A. Wittneben, "A new bandwidth efficient transmit antenna modulation diversity scheme for linear digital modulation," in *Proc. IEEE Int. Conf. Communications*, 1993, pp. 1630–1634.

[81] ——, "Base station modulation diversity for digital SIMULCAST," *Proc. IEEE Vehicular Technology Conf.*, pp. 505–511, May 1993.

[82] J. G. Proakis, *Digital Communications*, 3 ed. New York: McGraw-Hill, 1995.

[83] 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. 44, pp. 909–926, May 1998.

[84] P. Hoeher and J. Lodge, "Turbo DPSK: Iterative differential PSK demodulation and channel decoding," *IEEE Trans. Commun.*, vol. 47, pp. 837–843, June 1999.

[85] M. Peleg, I. Sason, S. Shamai (Shitz), and A. Elia, "On interleaved, differentially encoded convolutional codes," *IEEE Trans. Inform. Theory*, vol. 45, pp. 2572–2582, Nov. 1999.

[86] L. R. 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.

[87] S. ten Brink, "Convergence of iterative decoding," *IEE Electron. Lett.*, vol. 35, no. 10, pp. 806–8, May 1999.

[88] ——, "Convergence behavior of iteratively decoded parallel concatenated codes," *IEEE Trans. Commun.*, submitted for publication.

[89] ——, "A rate one-half code for approaching the Shannon limit by 0.1 dB," *IEE Electron. Lett.*, vol. 36, no. 15, pp. 1293–1294, July 2000.

[90] ——, "Design of serially concatenated codes based on iterative decoding convergence," in *Proc. 2nd Int. Symp. Turbo Codes and Related Topics*, Brest, France, 2000.

[91] A. Grant, "Convergence analysis of nonbinary iterative decoding," in *Proc. IEEE Global Communications Conf.*, San Antonio, TX, Nov. 25–29, 2001.

[92] P. Alexander and A. Grant, "Iterative sequence estimation using unconstrained state-space," in *Proc. IEEE Int. Symp. Information Theory*, Boston, MA, Aug. 1998, p. 242.

[93] ——, "Iterative decoding and channel estimation," in *Proc. IEEE Int. Symp. Information Theory*, Sorrento, Italy, June 2000, p. 171.

[94] ——, "Iterative channel and information sequence estimation in CDMA," in *Proc. IEEE Int. Symp. Spread Spectrum Techniques and Applications*, Parsippany, NJ, Sept. 2000, Invited Paper, pp. 593–597.