Joint Decoding and Channel Estimation for Space-Time Codes

Alex Grant*
Institute for Telecommunications Research
University of South Australia
Mawson Lakes SA 5095, AUSTRALIA
`Alex.Grant@unisa.edu.au`

Abstract

We consider the problem of joint decoding and channel estimation for space-time trellis codes transmitted over flat slowly varying Rayleigh fading channels. We assume that the receiver has no a-priori knowledge of the channel parameters. The amount of training data required is reduced by using a decision feedback approach to iteratively improve channel estimates.

1. Introduction

Information theory has promised that coordinated coding over multiple transmit antennas, coupled with the use of multiple receive antennas, will yield a great increase in capacity for wireless fading channels [18, 5]. Frequently cited is the linear increase of capacity with the simultaneous increase in the number of transmit and receive antennas.

Code design for such channels has mostly concentrated on the case when the fading channel parameters are perfectly known at the receiver. These codes are known as space-time codes, introduced in [2, 17, 15]. In practice, performance depends upon the quality of the channel estimates available [16]. The very motivation for using space-time codes (high spectral efficiency) precludes the use of long training sequences.

More recently, both information theorists and code designers have turned their interest to the problem of unknown channel parameters. In [13], the capacity of the flat fading multiple antenna channel is found, and some interesting relationships between the number of antennas and the coherence time of the channel are found. Further results have been found, based on treating code symbols as *subspaces*, resulting in a sphere-packing problem on the Grassman manifold [19, 1]. Several authors have already attempted code design for these types of channels [12, 11, 9], in which no explicit use is made of training sequences. These code designs are in the spirit of *blind* techniques which use knowledge of the *structure* (in this case specially built into the code itself) of the transmitted signals in order to obtain data estimates. Channel estimates may or may not be calculated as an intermediate step in this process.

In [10], an information theoretic approach is used to determine how much training is required for the multiple antenna channel. In that paper, the authors assume MMSE channel estimation based on a known training sequence. The results of that channel estimation is then used for decoding of the unknown data portion of the transmission. This conveniently partitions the problem into a channel estimation step, followed by a data detection conditioned on the estimated data (treating it as if it were in fact exact). This is also the common engineering solution to the problem. In practice the channel state information (CSI) is never known *a-priori* by either the receiver or transmitter. In fact one way of viewing the use of a known training sequence is that it is a simple construction (with simple "decoding" strategy) for a code in which the channel is unknown.

The two basic approaches to the problem just described could be summarized as

a. Special code design (no explicit CSI estimated)

b. Separate training and decoding (training used to estimate CSI).

In this paper, we take a slightly different approach. We propose a receiver that uses only a small amount of training, coupled with an iterative process that uses both the explicit training data and the coded information to estimate the CSI. The training is used to provide preliminary channel estimates which are used as side information for decoding the unknown data signal. The redundancy present in the code provides some robustness to these poor channel estimates. Data estimates obtained via decoding are then used as uncertain training sequences for the purposes of improved channel estimation. This process is continued

---

[*] This work was supported by Nortel Networks.

in an iterative fashion until some stopping criteria is met. This can be viewed as either joint decoding/channel estimation, or (since we actually don't care about the accuracy of the CSI) sub-optimal decoding of a non-coherent code, in which CSI is regarded as a state variable.

This paper is organized as follows. In Section 2 we give a discrete time mathematical channel model. The proposed receiver is then described in Section 3. In Section 4 we apply the receiver to the problem of decoding space-time codes. We present simulation results which show that the receiver can approach the performance of the optimal decoder with perfect channel knowledge. This work extends the ideas of [4]. Related work may be found in [3], where joint decoding and channel estimation of multiuser frequency selective fading channels is considered. Iterative multiuser decoding has been considered in [8, 14]. Iterative decoding for flat fading multiple access channels is considered in [6]. The ideas of [4] have also been applied to decoding of space-time codes in [7], but they do not consider channel estimation.

We shall use the following notations. The vector $\mathbf{x} \in \mathbb{C}^n$ is an column vector with complex entries x_i, $i = 1, 2, \ldots, n$. Likewise $\mathbf{A} \in \mathbb{C}^{m \times n}$ is a matrix with complex entries A_{ij}, $i = 1, \ldots, m$, $j = 1, \ldots, n$. The superscript * denotes Hermitian adjoint for matrices and vectors, and complex conjugate for scalars. $\mathbf{I}_n$ is the $n \times n$ identity matrix. For a random variable X, $\mathsf{E}\left[X\right]$ is its expectation.

2. Channel Model

Specifically, we are interested in a communication channel with t transmit antennas and r receive antennas. Associated with each transmit/receive pair is a possibly time varying complex scalar channel gain (extension to inter-symbol interference channels is straightforward. In that case we consider impulse responses between each input and output). For simplicity we consider a discrete-time channel.

With reference to Figure 1, at each symbol interval (for clarity the time index is dropped on the Figure) $l = 1, 2, \ldots, L$, the received matched-filtered vector $\mathbf{y}[l] \in \mathbb{C}^r$ depends on the transmitted vector, $\mathbf{x}[l] \in \mathbb{C}^t$ according to

$$\mathbf{y}[l] = \mathbf{H}[l]\mathbf{x}[l] + \mathbf{n}[l]. \tag{1}$$

Element $y_j[l]$ is matched-filter output j, while $x_i[l]$ is the transmit signal at input i. Throughout this paper we shall assume that $\mathbf{x}[l]$ is a coded sequence of vectors, and we refer to these vector symbols as space-time symbols. The matrix $\mathbf{H}[l] \in \mathbb{C}^{r \times t}$ has as elements $H_{ji}[l] \in \mathbb{C}$, which are the complex channel gains (representing flat Rayleigh fading) between input i and output j at time l. For simplicity we may consider the elements of $H[l]$ as independent, although channel correlation is not precluded by our model. The vector $\mathbf{n}[l]$ contains i.i.d. circularly symmetric Gaussian noise samples, $\mathsf{E}\left[\mathbf{n}[l]\mathbf{n}^*[l]\right] = \sigma^2\mathbf{I}_r$. We place the following power constraint on the transmitted signal (independent of t), $\mathsf{E}\left[\mathbf{x}[l]^*\mathbf{x}[l]\right] \le P$. We denote the signal to noise ratio (SNR) as $\gamma = P/\sigma^2$.

This linear model can also be used to describe systems such as linear multiple-access and orthogonal frequency division modulation.

Flat Rayleigh fading can be modeled by (1) by letting $Hji[l]$ be complex Gaussian. If the fading is assumed to be uncorrelated over antennas (sufficient antenna spacing) then the $H_{ji}[l]$ are independent and circularly symmetric.

$$\mathsf{E}\left[H_{ji}[l]H^*_{j'i'}[l]\right] = \begin{cases} 1 & i = i', j = j' \\ 0 & \text{otherwise.} \end{cases}$$

For the purposes of this paper, we shall assume such independent fading, although our proposed technique can be easily extended to incorporate dependent fading. We do allow however that each channel parameter may be correlated in time.

3. Iterative Receiver

The receiver operates in a semi-blind manner, in that a small amount of training is used to start the iteration. We assume that the transmitted coded sequence $\mathbf{x}[l]$ corresponds to the space-time encoding of a concatenation of a sequence that is known to the receiver and an information sequence, unknown to the receiver. Encoding of the known sequence produces a coded training sequence, which (under the assumption that the receiver knows the initial encoder state) is also known to the receiver. We also assume that $\mathsf{E}\left[\hat{x}_i^*[l]x_i[l]\right] = 1$ (which is the case for phase shift keying).

Figure 2 shows the structure of the receiver. After the matched filter front end, a Viterbi algorithm finds the maximum likelihood sequence given the current channel estimate. At the first iteration, this channel estimate is based only upon the coded training symbols that were transmitted.

Let $\hat{x}_i[l]$ be the sequence of estimated code symbols for transmit antenna i and let $\mathbf{h}_i[l] \in \mathbb{C}^r$ be column i of $\mathbf{H}[l]$. The channel estimator then sequentially re-computes an "unconstrained" channel estimate according to

$$\hat{\mathbf{h}}_i[l] = \hat{x}_i^*[l]\left(\mathbf{y}[l] - \sum_{m \neq i} \hat{x}_m[l]\hat{\mathbf{h}}_m[l]\right) \tag{2}$$

This may be done either in parallel, i.e. columns of $\hat{\mathbf{H}}[l]$ are updated simultaneously, or serially, in which updated columns are used as soon as they are available for estimation of other columns. Note that in the absence of data or

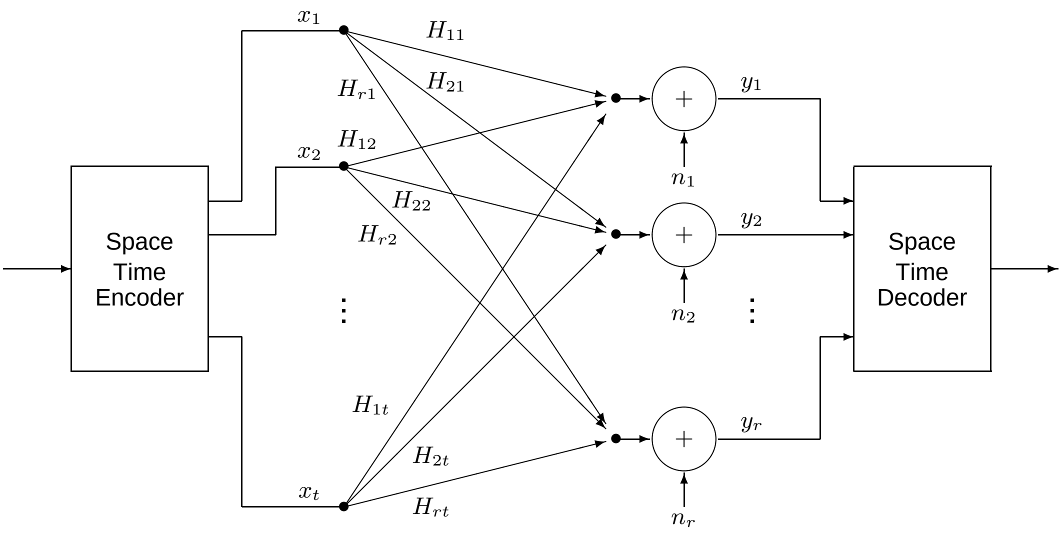

Figure 1. Multi-antenna system with flat fading.

channel estimation errors from the other users (2) yields $\hat{\mathbf{h}}_i[l] = \hat{x}_i^*[l]x_i[l]\mathbf{h}_i[l] + \hat{x}_i^*[l]\mathbf{n}[l]$.

Thus we attempt to estimate the channel coefficients between transmit antenna i and each of the receive antennas by first cancelling the current estimates of the signals received from the other transmit antennas.

Spectral constraints (e.g. known time correlations) on $\mathbf{h}$ are then enforced via MMSE filtering. As a simple example, if it is assumed that $\mathbf{H}[l] = \mathbf{H}$ is constant over some known coherence time L, we find

$$\hat{\mathbf{h}}_i = \frac{1}{\|\hat{\mathbf{x}}_i\|_1} \sum_{l=1}^{L} \hat{\mathbf{h}}_i[l].$$

We can show that this estimator is particularly robust to errors in the $\hat{\mathbf{x}}_i$, typically present in the first few iterations.

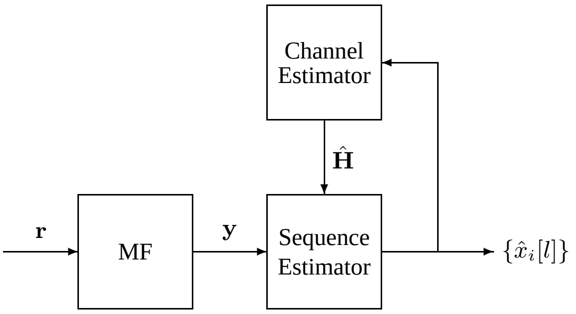

Figure 2. Joint decoder/channel estimator iterative receiver structure.

4. Simulation Results

Our simulations are based on the following situation. It is assumed that the channel is slowly varying such that $\mathbf{H}$ remains constant over each data frame, but is selected independently for each frame. Flat Rayleigh fading is modeled by choosing the H_{ji} with mutually independent Normal $N(0, 1/2)$ real and imaginary parts. Frame lengths of 130 space-time symbols are used (as in [17]).

We use joint Viterbi decoding of $\mathbf{x}$ treating the channel estimate from each iteration as if it were exact, i.e. we use the Euclidean branch metric

$$\left\| \mathbf{y}[l] - \hat{\mathbf{H}}\hat{\mathbf{x}}[l] \right\|_2 .$$

Note that if we can track the performance of the channel estimator, the decoding process may be improved by using the modified metric [16, Eqn. (4)]. In particular this requires knowledge of the error variance on the channel estimates (which are assumed normally distributed).

In order to provide a reference point for our results, Figure 3 shows the simulated performance of a conventional receiver, using dedicated orthogonal pilot transmissions for MMSE channel estimation. The four state space-time trellis code from [17, Fig. 4] is used, with $t = 2, r = 1$ and frame length 130 space-time symbols at 2 bits/sec/Hz.

For quasi-static fading the parameter of interest is the frame erasure rate (FER, square markers). Also shown is the bit error rate (BER, circle markers). Shown for reference are the simulated FER and BER performance for perfect CSI knowledge at the receiver (solid lines).

Figure 4 shows the simulated performance of the proposed receiver for the same parameters as for the previous simulation. Instead of orthogonal pilot sequences however, two coded training symbols (space-time symbols) per frame are transmitted to aid in initial channel estimation (equivalent to transmitting a pilot at -18dB with respect to the data). The corresponding (known to the receiver) input bits are chosen such that the training symbols are orthogonal, ensuring channel identifiability. The initial channel estimation based on these training symbols is maximum likelihood.

Perfect channel knowledge FER performance (solid line) is achieved after 2 iterations. BER comes within 1dB of perfect (solid line) and improves upon further iteration.

FER converges faster than BER, since further iteration of the receiver gradually removes errors in packets that are already counted as erased. Thus the receiver correctly decodes most packets that are correctly decoded under perfect channel knowledge, but suffers a higher BER on the remaining errored packets.

Figure 5 shows the performance results, now for $t = r = 2$ and the eight state code from [17, Fig. 5a]. Once again two coded training symbols were used. From these results we see that although the receiver is now operating at a lower SNR, we can still achieve the same FER as for perfect CSI, although now it takes four iterations.

Figure 5 shows performance results for the eight state code and $t = r = 2$. This time the number of training symbols is increased to three. This has the effect of speeding up the iteration slightly (perfect FER result obtained after 3 iterations). It also improves the BER performance. Thus we see that a simple trade-off can be made between complexity and spectral efficiency.

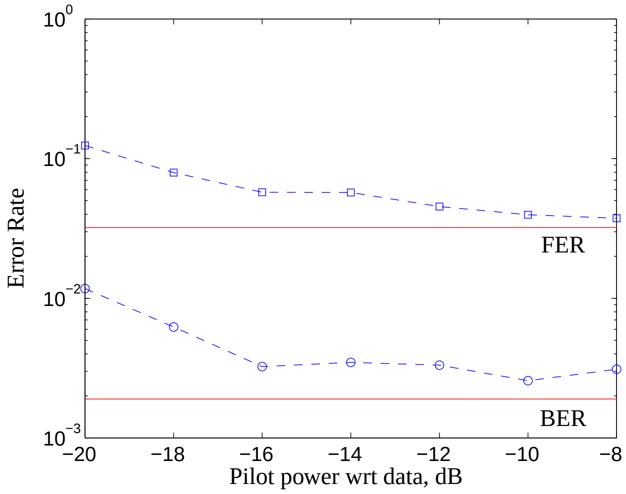

Figure 3. Conventional receiver performance results. $t = 2$**,** $r = 1$**, four state code.**

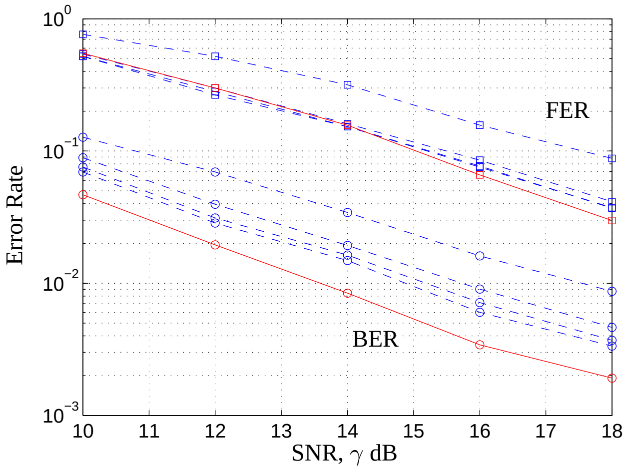

Figure 4. Performance results. $t = 2$**,** $r = 1$**, four state code.**

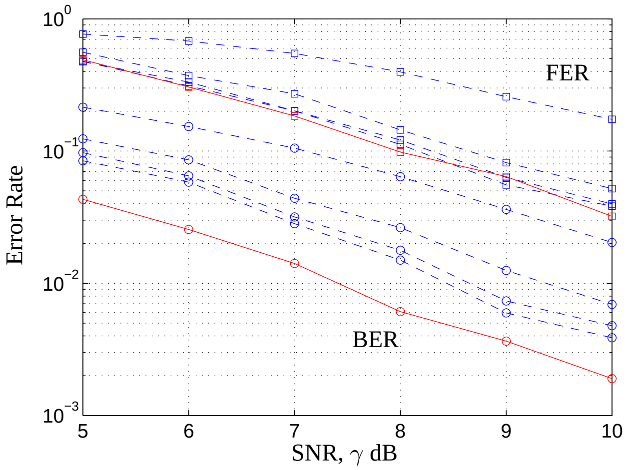

Figure 5. Performance results. $t = 2$**,** $r = 2$**, eight state code.**

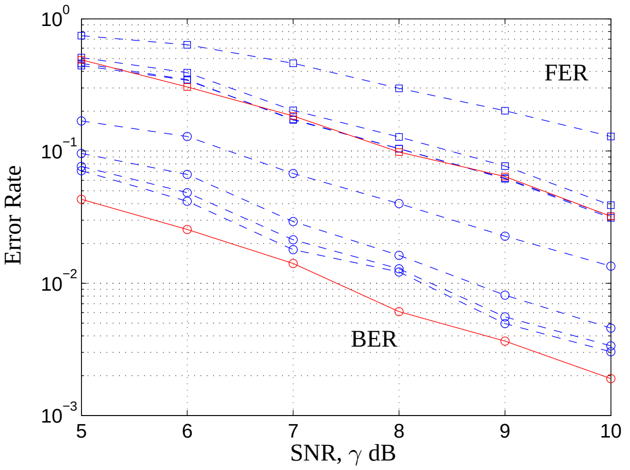

Figure 6. Performance results. $t = 2$**,** $r = 2$**, eight state code.**

5. Conclusion

We have investigated a simple technique for the joint detection and channel estimation of linear multiple-input multiple-output channels. Simulation results show that the receiver can approach the performance of a receiver with perfect channel knowledge using only very limited training data.

References

[1] D. Agrawal, T. J. Richardson, and R. Urbanke. Multiple-antenna signal constellations for fading channels. In *IEEE Int. Symp. Inform. Theory*, Sorrento, Italy, June 2000. Also submitted to IEEE Trans. Inform. Theory. Pre-print available at `http://degas.eecs.berkeley.edu/~dtse/pub.html`.

[2] S. M. Alamouti. A simple transmit diversity technique for wireless communications. *IEEE J. Selected Areas Commun.*, 16(8):1451–1458, Oct. 1998.

[3] P. Alexander and A. Grant. Iterative decoding and channel estimation. In *IEEE 2000 Int. Symp. Inform. Theory*, page 171, Sorrento, Italy, June 2000.

[4] P. D. Alexander, A. J. Grant, and M. C. Reed. Iterative detection on code-division multiple-access with error control coding. *European Transactions on Telecommunications*, 9(5):419–426, Sept.-Oct. 1998.

[5] G. J. Foschini and M. J. Gans. On limits of wireless communications in a fading environment when using multiple antennas. *Wireless Personal Communications*, 6(3):311–335, March 1998.

[6] H. E. Gamal and E. Geraniotis. Iterative multiuser detection for coded CDMA signals in AWGN and fading channels. *IEEE J. Selected Areas Commun.*, 18(1):30–41, Jan. 2000.

[7] H. E. Gamal and A. R. Hammons. The layered space-time architechture: A new perspective. *IEEE Trans. Inform. Theory*, 2000. Submitted.

[8] T. R. Giallorenzi and S. G. Wilson. Suboptimum multiuser receivers for convolutionally coded asynchronous DS-CDMA systems. *IEEE Trans. Commun.*, 44(9):1183–1196, Sept. 1996.

[9] B. Hassibi, B. Hochwald, A. Shokrollahi, and W. Sweldens. Multiple antennas and representation theory. In *IEEE Int. Symp. Inform. Theory*, Sorrento, Italy, June 2000. Also submitted to IEEE Trans. Inform. Theory. Pre-print available at `http://mars.bell-labs.com`.

[10] B. Hassibi and B. M. Hochwald. How much training is needed in multiple-antenna wireless links? Technical memorandum, Bell Laboratories, Lucent Technologies, April 2000. Available at `http://mars.bell-labs.com`.

[11] B. Hughes. Extensions to the theory of differential space-time modulation. In *IEEE Int. Symp. Inform. Theory*, Sorrento, Italy, June 2000. Also submitted to IEEE Trans. Inform. Theory.

[12] T. L. Marzetta, B. Hochwald, and B. Hassibi. Space-time autocoding: Arbitrarily reliable communication in a single fading interval. In *IEEE Int. Symp. Inform. Theory*, Sorrento, Italy, June 2000. Also submitted to IEEE Trans. Inform. Theory. Pre-print available at `http://mars.bell-labs.com`.

[13] T. L. Marzetta and B. M. Hochwald. Capacity of a mobile multiple-antenna communication link in Rayleigh flat fading. *IEEE Trans. Inform. Theory*, 45(1):139–157, Jan. 1999.

[14] M. Moher. An iterative multiuser decoder for near-capacity communications. *IEEE Trans. Commun.*, 46(7):870–880, July 1998.

[15] V. Tarokh, H. Jafarkhani, and A. R. Calderbank. Space-time block codes from orthogonal designs. *IEEE Trans. Inform. Theory*, 45(5):1456–1467, July 1999.

[16] V. Tarokh, A. Naguib, N. Seshadri, and A. R. Calderbank. Space-time codes for high data rate wireless communication: Performance criteria in the presence of channel estimation errors, mobility and multiple paths. *IEEE Trans. Commun.*, 47(2):199–207, Feb. 1999.

[17] 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*, 44:744–765, March 1998.

[18] I. E. Telatar. Capacity of multi-antenna Gaussian channels. *European Trans. Telecomm.*, 10(6):585–595, Nov.-Dec. 1999. Origninally published as Tech. Memo., Bell Laboratories, Lucent Technologies, October 1995.

[19] L. Zheng and D. Tse. Sphere packing in the Grassman manifold: a geometric approach to the noncoherent multi-antenna channel. In *IEEE Int. Symp. Inform. Theory*, Sorrento, Italy, June 2000. Also submitted to IEEE Trans. Inform. Theory.