Joint Decoding and Channel Estimation for Linear MIMO Channels

Alex Grant
Institute for Telecommunications Research
University of South Australia
Mawson Lakes Blvd, Mawson Lakes SA 5095
AUSTRALIA

Abstract— **We consider the problem of joint decoding and channel estimation for linear multiple-input multiple-output 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. We present experimental results for joint channel estimation and decoding of space-time channels.**

I. Introduction

Linear multiple input multiple output (MIMO) channels occur frequently in communications systems. Such channels are of interest whenever simultaneously transmitted signals are non-orthogonal, or are distorted by the channel in such a way that they are correlated at the receiver. The different channel inputs may be statistically independent (for example the various users of a multiple access system), or dependent (for example an OFDM system with inter-carrier interference or a spatial diversity channel).

In either case, the optimal receiver will generally perform joint detection of the transmitted signals given the received signals (based on whatever knowledge of the channel statistics that can be considered known). System performance can usually be improved if the detection is also conditional upon the channel state information. Such information is however usually not available. The optimal joint detector is frequently approximated through use of a separate channel estimator, conveniently partitioning 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).

In order for the receiver to reliably estimate the channel state information, a period of training is sometimes used, in which signals (training) known to the receiver are transmitted. Alternatively, the receiver may operate blind. Such receivers use knowledge of the *structure* of the transmitted signals in order to obtain channel estimates.

In this paper we investigate a simple sub-optimal method for partially blind reception of coded signals transmitted over MIMO channels. A minimal amount of 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 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 in an iterative fashion until some stopping criteria is met.

This work was supported by Nortel Networks.

Our contribution is a low complexity joint decoder/channel estimator for such systems. In Section II we give a discrete time mathematical channel model for the MIMO system. The proposed receiver is then described in Section III. In Section IV 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 [1] for unknown MIMO channels. Related work may be found in [2], where joint decoding and channel estimation of multiuser frequency selective fading channels is considered. Iterative multiuser decoding has been considered in [3], [4]. Iterative decoding for flat fading multiple access channels is considered in [5]. The ideas of [1] have also been applied to decoding of space-time codes in [6], 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. $\mathbf{I}_n$ is the $n \times n$ identity matrix. For a random variable X, $\mathsf{E}[X]$ is its expectation.

II. Channel Model

Specifically, we are interested in a communication channel with t inputs and r outputs. Associated with each input output 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 symbol synchronous discrete-time channel. In certain scenarios synchronicity is a valid assumption (such as single user space-time channels, or the down-link of a code-division multiple access system under a flat fading assumption).

At each symbol interval $l = 1, 2, \ldots, L$, 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

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

Element $y_j[l]$ is matched-filter output j, while $x_i[l]$ is the transmit signal at input i. 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 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, $\mathsf{E}[\mathbf{n}[l]\mathbf{n}^*[l]] = \sigma^2\mathbf{I}_r$. We place the following power constraint on the transmitted signal (independent of t), $\mathsf{E}\left[\mathbf{x}[l]^*\mathbf{x}[l]\right] \leq P$. We denote the signal to noise ratio (SNR) as $\gamma = P/\sigma^2$.

This linear model describes various scenarios of interest.

Linear multi-access. (E.g. DS-CDMA) in this case the elements $x_i[l]$ are independent and represent the transmissions of t independent users. The matrix $\mathbf{H}[l]$ describes the accessing method. For DS-CDMA column i of $\mathbf{H}[l]$ is the spreading sequence for user i at symbol interval l.

Orthogonal frequency division modulation. The elements of $\mathbf{x}[l]$ are the sub-carrier signals and $\mathbf{H}[l]$ models the inter-carrier interference due to channel imperfections.

Space diversity. In this case the elements of $\mathbf{x}[l]$ may be correlated. The matrix $\mathbf{H}[l]$ contains the possibly correlated path gains representing flat Rayleigh fading between transmit antenna i and receive antenna j. Such a model can represent both beam-forming type applications as well as space-time coded links.

We are interested in the general case in which $\mathbf{x}[l]$ is unknown, but obeys a known constraint (e.g. is convolutionally coded, or has passed through a known ISI channel) and the sequence of matrices $\mathbf{H}[l]$ is unknown, but may satisfy some known constraints. By considering some of the inputs $x[l]$ to be correlated fading processes [2] extends [1] in a different direction to obtain low complexity receivers for multiple-access channels with multi-path fading. In that work however the channel matrix $\mathbf{H}[l]$ (which represents accessing method of the users) is considered known.

III. Iterative Receiver

Figure 1 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 any training symbols that were transmitted. The decoder may operate either jointly (if complexity allows), or may be decomposed as to operate separately on the sequences $x_i[l]$. In this latter case the output sequence is no longer conditional ML.

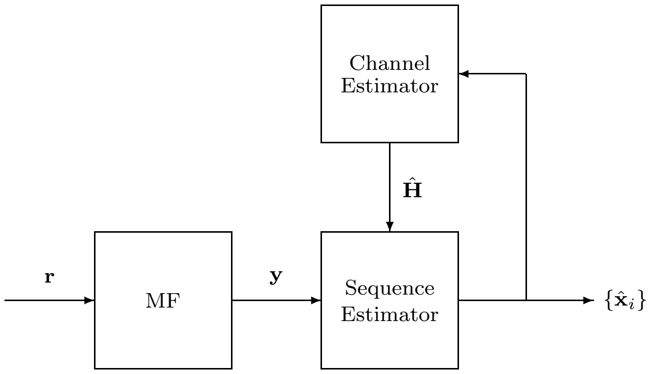

Fig. 1. Joint decoder/channel estimator terative receiver structure.

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

$$\hat{\mathbf{h}_i}[l] = \operatorname{conj}\left(\hat{x_i}[l]\right)\left(\mathbf{y}[l] - \sum_{m\neq i} \hat{x_m}[l]\hat{\mathbf{h}_m}[l]\right) \quad (2)$$

where conj denotes element-wise conjugation. 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 channel estimation errors from the other users (2) yields $\operatorname{conj}\left(\hat{x_i}[l]\right)x_i[l]\mathbf{h}_i[l]$. We assume that $\mathsf{E}\left[\operatorname{conj}\left(\hat{x_i}[l]\right)x_i[l]\right] = 1$. 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 on $\mathbf{h}$ are then enforced via MMSE filtering. For example if it is given 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.

IV. Space-Time Codes

Information theory indicates that improvements in capacity may be obtained on fading channels if joint encoding and decoding in the time and spatial dimensions is used [7].

Space-time codes [8] attempt to achieve these high data rates by jointly coding over multiple transmit antennas.

In the design of such codes it is usually assumed that perfect channel knowledge is available at the receiver. In practice, performance depends upon the quality of the channel estimates available [9]. The very motivation for using space-time codes (high data rates) precludes the use of long training sequences. It is therefore interesting to test the proposed receiver for decoding of space-time codes.

Figure 2 shows indicative performance of the proposed receiver for the following parameters: 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.

The four state code space-time trellis code from [8, Fig. 4] is used, with $t = 2, r = 1$ and frame length 384 bits (corresponding to 384 space-time symbol transmissions at 2 bits/sec/Hz).

Two coded training symbols (space-time symbols) per frame are transmitted to aid in initial channel estimation (equivalent to transmitting a pilot at -22dB 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.

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 [9, Eqn. (4)]. In particular this requires knowledge of the error variance on the channel estimates (which are assumed normally distributed).

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). 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.

V. Performance Analysis

Analysis of this receiver structure is possible using the idea of an iterated function system on some quality measure, e.g. noise variance at the output of the sequence estimator. Given some independence assumptions, we can show that contractiveness of both the sequence estimator and the channel estimator results in convergence to a fixed point, which is close to the perfect channel knowledge performance. Decoder failures are generally not due to this fixed point gradually moving away from the ideal, but rather the appearance of a second, usually unstable fixed point (a function of "system load", t/r). At the appearance of the unstable fixed point, system performance deteriorates suddenly and the system fails. Details of this analytical technique may be found in a forthcoming paper.

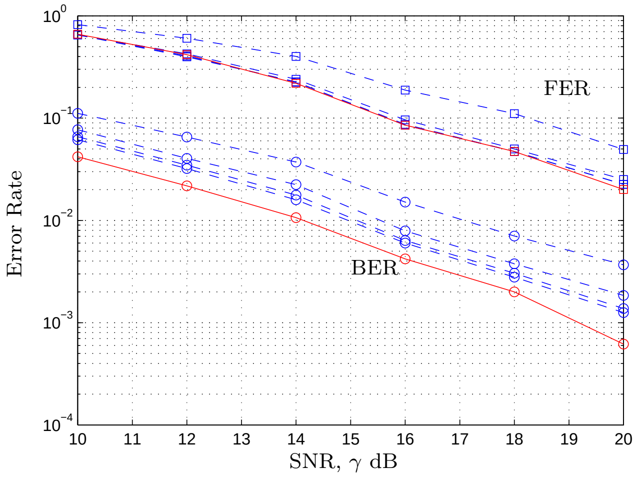

Fig. 2. Indicative performance results for the proposed iterative receiver.

VI. 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.

Acknowledgments

Thanks to P. Alexander for his ongoing collaborations in this area.

References

[1] 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*, vol. 9, no. 5, pp. 419–426, Sept.-Oct. 1998.

[2] P. Alexander and A. Grant, "Iterative decoding and channel estimation," in *IEEE 2000 Int. Symp. Inform. Theory*, Sorrento, Italy, June 2000, To Appear.

[3] T. R. Giallorenzi and S. G. Wilson, "Suboptimum multiuser receivers for convolutionally coded asynchronous DS-CDMA systems," *IEEE Trans. Commun.*, vol. 44, no. 9, pp. 1183–1196, Sept. 1996.

[4] Michael Moher, "An iterative multiuser decoder for near-capacity communications," *IEEE Trans. Commun.*, vol. 46, no. 7, pp. 870–880, July 1998.

[5] H. El Gamal and E. Geraniotis, "Iterative multiuser detection for coded CDMA signals in AWGN and fading channels," *IEEE J. Selected Areas Commun.*, vol. 18, no. 1, pp. 30–41, jan 2000.

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

[7] İ. Emre Telatar, "Capacity of multi-antenna Gaussian channels," *European Trans. Telecomm.*, vol. 10, no. 6, Nov.-Dec. 1999.

[8] Vahid Tarokh, Nambi 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, March 1998.

[9] Vahid Tarokh, Ayman Naguib, Nambi 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.*, vol. 47, no. 2, pp. 199–207, Feb. 1999.