Pilot Embedding for Joint Channel Estimation and Data Detection in MIMO Communication Systems

Haidong Zhu, Behrouz Farhang-Boroujeny, *Senior Member, IEEE*, and Christian Schlegel, *Senior Member, IEEE*

***Abstract*—A novel pilot-aided joint channel estimation and data detection method for MIMO communication systems is proposed. Unlike, conventional methods where pilots are time-multiplexed with data symbols, a pilot-embedding method where low-level pilots are transmitted concurrently with the data is used to obtain an initial estimate of the channel such that a turbo decoding process can be started. The soft information obtained from the turbo decoder is subsequently used to improve channel estimates.**

***Index Terms*—Channel estimation, MIMO channels, multiple antenna systems, turbo decoding.**

I. Introduction

THE potential capacity of multiple-input multiple-output (MIMO) channels has recently been recognized [1], [2] and a widespread effort to harness the capacity of such channels has began. Clever channel estimation techniques and capacity achieving data detection methods are the essential components of this endeavor.

Efforts to develop receiver structures that identify the channel while detecting data have been taken up recently. In [3], for example, it is shown that turbo coding principles can be combined with a simple channel estimator to construct systems that come close to the capacity limit. The system begins with a coarse estimate of the channel and improved estimates of the channel are later obtained by integrating the estimation of the channel into the decoding loop. Soft information from the iterative decoder is used to improve channel estimates after every iteration of the decoder. The results presented in [3] show that such joint decoding and channel estimation loses very little compared to an idealized system where perfect channel information is known. The joint decoding and channel estimation scheme that has been proposed in [3] relies on the differential property of the differential space-time (DST) codes. In the absence of a differential structure in the code, we need other methods of obtaining such estimates. In this letter, we propose one such method.

We propose transmission of pilot symbols that are transmitted along with data. These pilots, which are at a level much lower than signal level, are used to obtain the initial coarse estimates of the channel. We note that this scheme is different from the more conventional pilot aided methods, [4], where pilots are time-multiplexed with data and thus consume portion of the valuable system bandwidth. We will comment on advantages of the proposed pilot embedding when compared with its time-multiplexed counterpart in Section VI.

Manuscript received July 10, 2002. The associate editor coordinating the review of this letter and approving it for publication was Dr. J. Ritcey.

H. Zhu and B. Farhang-Boroujeny are with the Department of Electrical and Computer Engineering, University of Utah, Salt Lake City, UT 84112 USA (e-mail: farhang@ece.utah.edu).

C. Schlegel was with the Department of Electrical and Computer Engineering, University of Utah, Salt Lake City, UT 84112 USA. He is now with Department of Electrical and Computer Engineering, University of Alberta, Edmonton, AB T6G 2V4, Canada.

Digital Object Identifier 10.1109/LCOMM.2002.807434

II. MIMO System Model

In general, a ST symbol $\mathbf{C}_l$ is a $L_t \times L_c$ codeword matrix that is transmitted across all the transmit antennas in L_c time slots. The received symbol matrix collected at the receiver is then

$$\mathbf{Y}_l = \mathbf{H}_l \mathbf{C}_l + \mathbf{N}_l \tag{1}$$

where $\mathbf{H}_l$ is the $L_r \times L_t$ matrix of the channel, $\mathbf{N}_l$ is an $L_r \times L_c$ matrix of noise samples, and L_t and L_r are the numbers of transmit and receive antennas, respectively. Independent Rayleigh fading is modeled by selecting the elements of $\mathbf{H}_l$ as independent unit variance complex Gaussian random variables. We distinguish between *fast* fading, in which $\mathbf{H}_l$ evolves according to a process whose dominant frequency is much faster than $1/L_b$, where L_b is the data block (or, packet) length, but much slower than $1/L_c$ and *quasi-static* fading, in which $\mathbf{H}_l$ is selected independently and then held constant over each data block. Simulation scenarios that are examined in this letter are those of the fast fading channels.

III. Pilot Embedding for Channel Estimation

The idea of pilot embedding was first proposed in [6] in the context of single-input single-output single-carrier systems and has recently been extended to multicarrier systems [7]. The initial work in [6] uses pilot symbols to obtain an initial estimate of the channel so that subsequently the receiver can work in a decision directed mode for tracking. In [7] pilot symbols are used to initialize an iterative joint channel estimation and data detection loop in an OFDM system that uses a simple slicer detector. The present work extends the idea of pilot embedding to ST turbo coded systems. The goal is to demonstrate feasibility of this technique in systems that are designed to perform near channel capacity. The proposed method works as follows.

At the transmitter, the pilot matrix $\mathbf{P}$ is added to the codewords $\mathbf{C}_l$. To decorrelate the estimates of different columns of $\mathbf{H}_l$, a set of orthogonal vectors, such as Walsh-Hadamard codes, are assigned to different rows of $\mathbf{P}$. Since, in general, $\mathbf{P}$ may overlap with a number of codewords $\mathbf{C}_l$, we define $\mathbf{C}_l^p = [\ldots, \mathbf{C}_{l-1}, \mathbf{C}_l, \mathbf{C}_{l+1}, \ldots]$ of the same size as $\mathbf{P}$. The corresponding received signal matrix is thus given by

$$\mathbf{Y}_l^p = \mathbf{H}_l \left(\mathbf{C}_l^p + \mathbf{P}\right) + \mathbf{N}_l^p \tag{2}$$

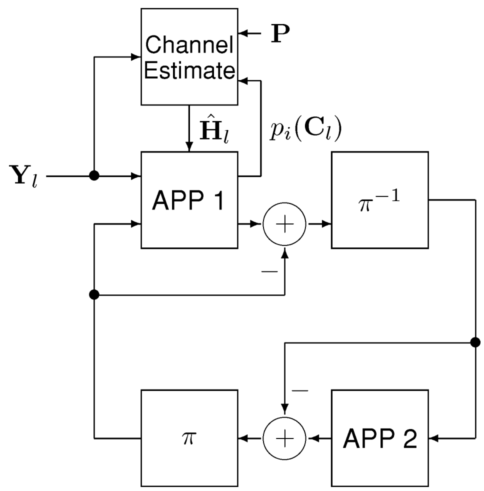

Fig. 1. Integration of channel estimator within a turbo decoder system. Pilots $\mathbf{P}$ are used to obtain a first estimate of the channel and to start the decoding process.

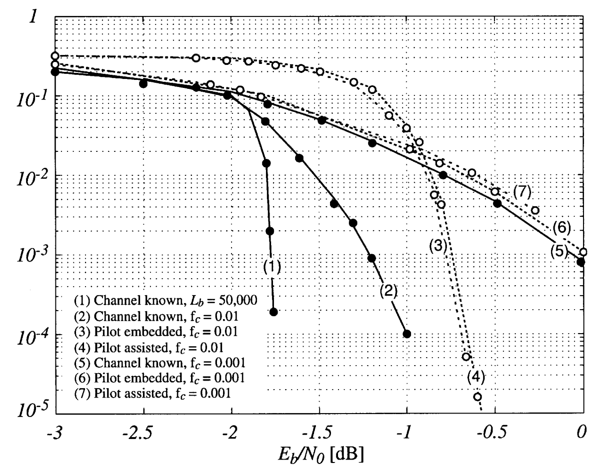

Fig. 2. Bit error rate results. E_b/N_0 is bit energy over noise power.

where $\mathbf{Y}_l^p$ and $\mathbf{N}_l^p$ are defined analogously to $\mathbf{C}_l^p$. Postmultiplying both sides of (2) by $\mathbf{P}^H$, using the fact that $\mathbf{P}\mathbf{P}^H = k\mathbf{I}$ and rearranging the result, we obtain

$$\hat{\mathbf{H}}_l = \frac{1}{k}\mathbf{Y}_l^p\mathbf{P}^H = \mathbf{H}_l + \mathbf{N}_l' \quad (3)$$

where $\mathbf{N}_l' = (1/k)(\mathbf{H}_l\mathbf{C}_l^p\mathbf{P}^H + \mathbf{N}_l^p\mathbf{P}^H)$ is the estimation error.

IV. Integration of Channel Estimation With Data Detection

Fig. 1 presents a schematic that shows how the channel estimation and data detection are integrated into a turbo decoding system. The receiver begins by using pilot symbols $\mathbf{P}$ to obtain a first estimate of the channel. The turbo decoder consisting of *a posteriori* decoder APP 1, deinterleaver π^{-1}, *a posteriori* decoder APP 2 and interleaver π obtains a-posteriori probabilities $p_i(\mathbf{C}_l) = \Pr(\mathbf{C}_l = \mathbf{C}^{(i)})$, where $\mathbf{C}^{(i)}$, for different values of i, are possible choices of the codeword. The channel estimator then uses $p_i(\mathbf{C}_l)$ to get a more accurate estimate of the channel for the next iteration of the decoder. Given $p_i(\mathbf{C}_l)$, a soft estimate of $\mathbf{C}_l$ is obtained as $\hat{\mathbf{C}}_l = \sum_i p_i(\mathbf{C}_l)\mathbf{C}^{(i)}$. These estimates are combined with the pilot matrix $\mathbf{P}$ to obtain an estimate of the transmitted signal as $\hat{\mathbf{C}}_l^p + \mathbf{P}$, where $\hat{\mathbf{C}}_l^p$ is defined in a similar way as $\mathbf{C}_l^p$. Accordingly, the least-squares/maximum-likelihood estimate of the channel is obtained as

$$\hat{\mathbf{H}}_l = \mathbf{Y}_l^p\left(\hat{\mathbf{C}}_l^p + \mathbf{P}\right)^H\left(\left(\hat{\mathbf{C}}_l^p + \mathbf{P}\right)\left(\hat{\mathbf{C}}_l^p + \mathbf{P}\right)^H\right)^{-1} \quad (4)$$

V. Computer Simulations

We consider a serially concatenated turbo coded communication system with $L_t = L_r = 2$. The outer code is a rate 2/3 standard maximum free distance convolutional code with 4 states [8], [9]. The inner code is a rate 3/4 trellis code that we have designed through an exhaustive search. There is also a random interleaver between the inner code and the transmit antennas. This interleaver serves to randomize/uncorrelate the channel gains at the adjacent codewords. The advantage of such additional interleaver in improving the system performance has also been noted in some previous works; see, for example, [10]. Each codeword is transmitted in two time slots, i.e., $L_c = 2$ and carries 2 bits. Hence, we effectively transmit 1 b/s/Hz. The details of this code (or any other code that may be used) are not important for the conclusion of this letter. However, the present code is chosen to demonstrate that the proposed scheme works well even when the system is designed to work near channel capacity. Such cases are usually characterized by high noise levels. Examples given later address cases where the SNR is negative.

We consider a time-varying channel which undergoes multiple fades within each block of the turbo code, which contains 5000 symbols. Pilot symbols at 10 dB below the signal level are added and used to initialize the estimates of the channel. A Rayleigh fading channel based on Jakes' model [5] is considered. Results for the normalized (to the symbol rate) fading rates of $f_c = 0.01$ and $f_c = 0.001$ are presented. The pilot matrix $\mathbf{P}$ that we have used has the simple structure of

$$\mathbf{P} = \alpha\begin{bmatrix} \cdots & 1 & 1 & 1 & 1 & \cdots \\ \cdots & 1 & -1 & 1 & -1 & \cdots \end{bmatrix}$$

where α is a constant which determines the pilot power level. The number of columns of $\mathbf{P}$ are chosen equal to 40 when $f_c = 0.01$ and equal to 80 when $f_c = 0.001$.

Fig. 2 presents the bit-error-rate (BER) results of the proposed pilot embedding method. For comparison, we have also shown the results of the pilot assisted method (discussed further in the next section) as well as for a perfectly known channel. We note that for $f_c = 0.001$, estimating the channel loses very little compared to the case of a perfectly known channel. For $f_c = 0.01$ the difference between the known and estimated channel is more significant. At $\mathrm{BER} = 10^{-3}$ this difference is about 0.5 dB, however, diminishes as the BER decreases. We note that, for the present comparison, the pilot power has been excluded in the computation of SNR. If included, it will result in a right shift of the BER curves by ($10\log 1.1 =$) 0.414 dB.

Also shown in Fig. 2 are the BER results corresponding to a very long block length of $L_b = 50{,}000$ symbols. This result serves to show the limit of the present code which could only be achieved when L_b is sufficiently long. For shorter code lengths (such as, $L_b = 5000$), an increase in BER is due to the fact that for some blocks the actual SNR falls significantly below the average SNR. This is particularly evident for $f_c = 0.001$. We also note that for 1 b/s/Hz, the Shannon bound is -3.1 dB.

VI. Embedded Pilots Versus Time-Multiplexed Pilots

As we noted earlier (in Section I), a more common approach to using pilots for channel estimation is to time-multiplex pilots with data symbols. Conventionally, pilot assisted schemes of this sort are used along with a channel interpolation based on Wiener filters to obtain a reliable estimate of the channel [4]. In the channel scenario that we consider here (characterized by a very low SNR and possibly a very fast fading rate) such reliable estimates could only be obtained if a relatively large number of pilot symbols could be used. On the other hand, we may recall that the joint channel estimation and data detection scheme that was presented above can begin with a very coarse estimate of the channel. Hence, when adopting pilot assisted scheme, it is also possible to begin with a coarse estimate of the channel and this may be obtained by using a small number of pilot symbols.

To compare the performance of the proposed pilot embedding and the conventional pilot assisted scheme, in Fig. 2, we have also presented the BER results of the latter when a pilot symbol is inserted after every ten data symbols. As observed, the two methods have virtually the same performance. We note that the total transmit power for each block of data in both systems is the same. However, the pilot embedding scheme transmits each block of data in a shorter duration of time. It may thus be argued that the pilot embedding is more bandwidth efficient than the pilot assisted scheme.

VII. Conclusion

We proposed embedded pilots to initialize a turbo decoder when the channel is unknown to the receiver. Computer simulations show that the proposed method works very well for MIMO channels and allows operation very close to the channel capacity. Furthermore, the conventional pilot inserted methods, [4], could also be used for encoder initialization. However, the pilot embedding was found to be more bandwidth efficient since pilot symbols are transmitted along with data.

References

[1] E. Telatar, "Capacity of multi-antenna Gaussian channels," *Eur. Trans. Telecommun.*, pp. 585–595, Nov.–Dec. 1999.

[2] G. J. Foschini and M. J. Gans, "On limits of wireless communications in a fading environment," *Wireless Commun. Mag.*, vol. 6, pp. 311–335, Mar. 1998.

[3] C. Schlegel and A. Grant, "Concatenated space-time coding," in *PIMRC 2001*, San Diego, CA, Sept. 30–Oct. 3.

[4] J. K. Cavers, "An analysis of pilot symbol assisted 16QAM for digital mobile communications," *IEEE Trans. Veh. Technol.*, vol. 40, pp. 686–693, Nov. 1991.

[5] W. C. Jakes, *Microwave Mobile Communications*: Wiley, 1974.

[6] B. Farhang-Boroujeny, "Pilot-based channel identification: A proposal for semi-blind identification of communication channels," *Electron. Lett.*, vol. 31, pp. 1044–1046, June 1995.

[7] F. Chin, B. Farhang-Boroujeny, and C. K. Ho, "Added pilot semi-blind channel estimation scheme for OFDM in fading channels," in *Proc. IEEE Globecom'01*, vol. 5, San Antonio, TX, Nov. 25–29, 2001, pp. 3075–3079.

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

[9] C. Schlegel, *Trellis Coding*. Piscataway, NJ: IEEE Press, 1997.

[10] A. Stefanov and T. M. 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.