Bit-interleaved coded modulations for multiple-input multiple-output channels

Joseph J. Boutros, Francesc Boixadera & Catherine Lamy
Communications and Electronics Department
Ecole Nationale Supérieure des Télécommunications
46 Rue Barrault 75013 Paris, France
e-mail: boutros@enst.fr

Abstract — **We describe a soft feedback iterative decoding technique for bit interleaved coded modulations (BICM) where the modulator uses multiple transmitting antennas and the receiver employs space-time coherent detection with multiple antennas. We investigate the bit error rate and the frame error rate performance of this technique applied to two families of error-control codes: simple nonrecursive binary convolutional codes and parallel turbo codes based on two recursive constituents. Two different types of fading channels are considered: independent flat Rayleigh and block fading channels. A 64-state rate 1/2 binary convolutional code performs 1.5 dB from outage capacity with perfect channel state information and 3 dB with Expectation-Maximization estimation on the block fading channel.**

I. Introduction

Iterative a posteriori probability (APP) decoding of compound codes is a very popular and an extremely efficient tool especially when a reasonable complexity maximum likelihood decoder is not available [1][2][3]. Iterative detection is not to be limited to compound error-correcting codes, since it can be applied to any concatenation whose elements are simple enough to make an easy evaluation of the *a posteriori* probability via the Forward-Backward algorithm [4] or more generally a soft-input soft-output (SISO) decoder [5]. This is the case for the structure of a bit interleaved coded modulation [6] [7] which can be considered as the serial concatenation of an outer error-correcting code and a channel modulator.

Multiple-input multiple-output (MIMO) channels permit a wide increase in the systems' spectral efficiency [8][9] and hence to reach modern technologies needs, as unveiled by Tarokh *et al.* for trellis codes [10] or more recently for turbo codes [11]. Combined with bit interleaved coded modulation and APP decoding, this technique is even more efficient. As shown later in this paper, using a 64-state convolutional code of rate 1/2 and a total spectral efficiency of 2 bits per symbol period, the obtained performance is only 1.5 dB away from outage capacity on a block fading channel.

The assumption we made according to which the channel state information is known at the receiver end is validated by our implementation of an Expectation-Maximization (EM) [12][13] channel estimator naturally included in the iterative decoding scheme. It is indeed shown that a loss of only 1.5 dB occurs when the receiver recovers information on the channel state on its own via EM.

The following section presents the system model and the notations. The derivation of coded bits a posteriori probability and the whole iterative procedure combining multiple antenna channel marginalization and SISO decoding are described in Section III. Finally, simulation results of classical non recursive binary convolutional codes and parallel turbo codes are presented in Section IV with two transmitting and two receiving antennas.

II. System model and notations

The wireless medium linking the n_t transmitting (Tx) antennas and the n_r receiving (Rx) antennas is assumed to be a flat fading multiple-input multiple-output channel. At time index $k \in \mathbb{Z}$, the channel output is the superposition of the n_t transmitted symbols weighted by the channel response. This can be expressed by

$$\mathbf{y}(k) = H(k) \cdot \mathbf{x}(k) + \mathbf{n}(k) \tag{1}$$

with $\mathbf{y}(k) = (y_i(k))^t{}_{i=1..n_r}$ the Rx signal vector, $\mathbf{x}(k) = (x_j(k))^t{}_{j=1..n_t}$ the Tx signal vector, $H(k) = [h_{i,j}(k)]_{i=1..n_r, j=1..n_t}$ the channel matrix and $\mathbf{n}(k) = (n_i(k))^t{}_{i=1..n_r}$ the additive white complex Gaussian noise with zero mean and variance $2\sigma^2 = 2N_0$ at time instant k.

The symbols x_j belong to a PSK or a QAM constellation of size $M = 2^m$. The fading coefficients $h_{i,j}(k) \in \mathbb{C}$ are Gaussian, mutually independent and satisfy $E[|h_{i,j}(k)|^2] = 1$. For clarity of notations the time index is omitted in the rest of the paper. The superscripts ^{t} and ^{h} stand for the transpose and the transpose conjugate respectively.

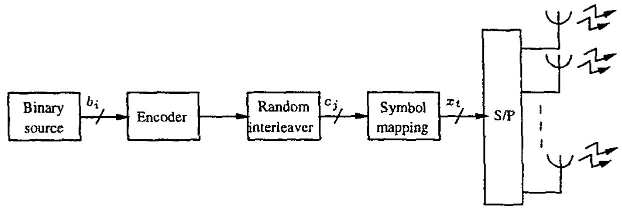

Figure 1: Multiple antennas bitwise transmitter.

The transmitter structure is illustrated in Fig. 1. The information bits $\mathbf{b} = (b_i)_{i=1}^{N_b}$ are encoded into N_c coded bits $\mathbf{c} = (c_j)_{j=1}^{N_c}$ which are then randomly interleaved and mapped into PSK/QAM symbols x_j. The block of N_c/m symbols to be transmitted is then divided into sub-blocks of length n_t and sent in parallel by the Tx antennas. At every time index, the signal vector $\mathbf{x}$ is hence a function of $m \times n_t$ coded bits

$$\mathbf{x} = (x_1 \cdots x_{n_t})^t = f(c_1, c_2, \cdots, c_{mn_t}) \tag{2}$$

In this paper we present two different encoder types: classical non-recursive non-systematic convolutional codes [14] and parallel turbo codes [1][2][15]. As we show later, the proposed system can be easily extended to other types of error control codes, provided that a SISO decoder is available. Furthermore, two types of channels are considered. The first type is a non static Rayleigh channel where the complex Gaussian distributed coefficients $h_{i,j}(k)$ change randomly and independently at each symbol period. The second type is a static block fading channel that maintains constant the coefficients $h_{i,j}(k)$ inside the same coded frame, $k = 1 \dots N_c/(mn_t)$.

III. Iterative detection and APP decoding

To recover the binary stream, soft information on the coded bits needs to be extracted from the received signals, or more precisely from the contribution of each transmitting antenna. Given the whole received signal over all time index $k = 1 \dots N_c/(mn_t)$, and since coded bits have been randomly interleaved, it is possible to compute the probability of c_j being equal to 0 or 1. This probability, called *a posteriori* probability $APP(c_j)$ can be expressed as

$$\begin{aligned} APP(c_j) &= p(c_j|\mathbf{y}) \\ &= \frac{p(\mathbf{y}|c_j) \cdot \pi(c_j)}{p(\mathbf{y})} \quad j = 1, \cdots, mn_t \\ APP(c_j) &\propto \pi(c_j) \cdot p(\mathbf{y}|c_j) = \pi(c_j) \cdot obs(c_j) \end{aligned} \tag{3}$$

where $\pi(c_j)$ is the *a priori* probability of the bit c_j and the observation $obs(c_j) = p(\mathbf{y}|c_j)$.

The conditional probability density $p(\mathbf{y}|c_j)$ is determined by marginalizing the joint density of all bits and the observation when taking into account that the received signals y_r are independent conditionally to the coded bits $c_1, \cdots, c_{mn_t}$

$$\begin{aligned} p(\mathbf{y}|c_j) &= \sum_{\substack{c_i \in \{0,1\} \\ i = 1 \cdots mn_t, i \neq j}} p(\mathbf{y}, c_1, .., c_{j-1}, c_{j+1}, .., c_{mn_t}|c_j) \\ &= \sum_{c_i \in \{0,1\}, i = 1 \cdots mn_t, i \neq j} p(\mathbf{y}|c_1, \cdots, c_{mn_t}) \prod_{\ell \neq j} \pi(c_\ell) \\ &= \sum_{\substack{c_i \in \{0,1\} \\ i = 1 \cdots mn_t, i \neq j}} \left(\prod_{r=1}^{n_r} p(y_r|c_1 \cdots c_{mn_t}) \prod_{\ell \neq j} \pi(c_\ell) \right) \end{aligned} \tag{4}$$

The conditional density $p(y_r|c_1, \cdots, c_{mn_t})$ is evaluated by

$$p(y_r|c_1, \cdots, c_{mn_t}) = \frac{e^{-\frac{\left\| y_r - \sum_{t=1}^{n_t} h_{t,r} x_t \right\|^2}{2\sigma^2}}}{(2\pi\sigma^2)} \tag{5}$$

where the x_t are defined by equation (2).

Notice that a less complex but sub-optimal method is also available to compute an approximation of the bit APPs. The received signals are independently processed at each Rx antenna to obtain several partial channel likelihoods. The total APP approximation is then

$$APP_{sub}(c_j) \propto \pi(c_j) \cdot \prod_{r=1}^{n_r} p(y_r|c_j) \tag{6}$$

where $p(y_r|c_j)$ is derived by marginalizing the total likelihood as in equation (4). Computer simulations showed that this second APP evaluation method is less optimal than the first one in terms of both binary error rate and convergence speed. Mathematically, this second method is sub-optimal because the received signals y_r are correlated when they are partially conditioned.

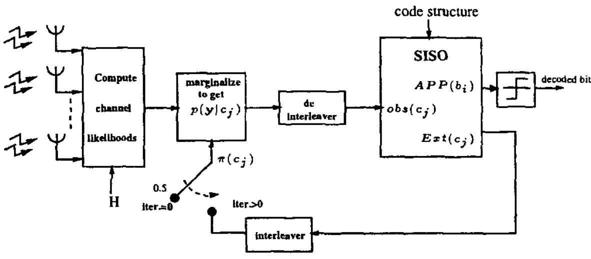

Figure 2: Multiple antennas bitwise receiver.

The evaluation of the conditional likelihood $p(\mathbf{y}|c_j) = obs(c_j)$ corresponds to the detection stage located at the receiver front. This likelihood $obs(c_j)$, called *observation* associated to c_j, is then processed by a soft-input soft-output decoder that takes into account the error correcting code constraints. The SISO decoder [5][16] generates an extrinsic information $Ext(c_j)$ which is equivalent to a new *a priori* probability $\pi(c_j)$ for the coded bit c_j. Hence, it is convenient to feed this *a priori* information back to the likelihood detector defined by equation (4). Thus, iterative detection and decoding is an excellent way to improve the estimation of the *a posteriori* probabilities.

Fig. 2 illustrates the iterative detection and decoding receiver structure. Note that the receiver is separated into two parts: the first part is non iterative and computes the received signal conditional probability at every antenna r according to formula (5). The second receiver part is iterative and its input depends also on the *a priori* probabilities. The final decision is made out of the *a posteriori* probability generated by the SISO decoder at the last iteration.

The whole iterative receiving scheme is performed as follows:

- Initialization: precompute $N_c/(mn_t) \times n_r$ likelihoods $p(y_r|c_1, \cdots, c_{mn_t})$ from the channel output. Set the N_c *a priori* probabilities $\pi(c_j)$ to $1/2$.
- Execution at each iteration: compute the N_c observations $p(\mathbf{y}|c_j)$ from the channel output using formula (4) where the *a priori* probabilities $\pi(c_\ell)$ are set to the extrinsic information $Ext(c_\ell)$ produced by the SISO decoder at the previous iteration. Then, apply the SISO decoder to compute $APP(c_j) \propto Ext(c_j) \times obs(c_j), j = 1 \cdots N_c$ and the *a posteriori* probabilities $APP(b_i), i = 1 \cdots N_b$.
- Final decision: at the last iteration, decide that $b_i = 0$ if $APP(b_i = 0) > APP(b_i = 1)$ otherwise $b_i = 1$.

The iterative procedure for MIMO APP detection presented above assumes that the channel parameters H and N_0 are known at the receiver end. Those parameters cannot be estimated in the case of a non-static Rayleigh fading channel. On the contrary, several methods can be employed for state estimation on the static block fading channel. The classical technique is the insertion of pilot symbols in the data frame, but it has the major drawback to reduce seriously the spectral efficiency of the system, especially when used at low signal-to-noise ratios. Another possible method is to perform the estimation with the Expectation-Maximization algorithm [12][17] with or without the help of pilot symbols as already done for estimating signal amplitude and noise variance in turbo decoding with BPSK modulation [18] and for continuous phase modulations (CPM) [19]. Iterative EM estimation formulas given below are proved in [13]

$$H^{i+1} = \sum_{k=1}^{\frac{N_c}{m n_t}} \mathbf{y}(k) E_{|H^i, N_0^i}[\mathbf{x}^h] \times \left(\sum_{k=1}^{\frac{N_c}{m n_t}} E_{|H^i, N_0^i}[\mathbf{x}\mathbf{x}^h] \right)^{-1}$$

$$N_0^{i+1} = \frac{1}{4 \frac{N_c}{m n_t}} \sum_{k=1}^{\frac{N_c}{m n_t}} E_{|H^i, N_0^i}[||\mathbf{y}(k) - H^{i+1}\mathbf{x}||^2] \qquad (7)$$

Numerical simulations showed that a completely blind initialization is not successful. Consequently, pilot symbols are used to derive the initial value (H^0, N_0^0) (e.g. 11% pilots as illustrated in Section IV). For subsequent iterations, the EM algorithm is naturally embedded in the detection-decoding iterations as it only needs the channel observations and *a posteriori* probabilities given by the SISO decoder. Thus, the added complexity is negligible.

IV. Results

Figures 3, 4 and 5 display the performance of a turbo code and a non-recursive non-systematic convolutional code over a non-static Rayleigh fading channel and a static block fading channel. The considered turbo code is a parallel concatenation of two recursive systematic binary convolutional codes with octal generators $(23, 35)$ and total frame length $N_c = 2000$ bits, interleaver size 1000. The convolutional code is binary non systematic non recursive with octal generators $(133, 171)$ and total frame length $N_c = 200$ bits. In all cases, the trellis of the $(133, 171)$ convolutional code and the trellis of the $(23, 35)$ turbo code constituent are correctly driven into the null state after the generation of N_c coded bits.

The gain achieved via iterative decoding is obvious on both considered channels and both codes. It has to be noted that this gain is obtained almost completely after three iterations only, which corresponds to a reasonable delay. Whereas on the independent Rayleigh fading channel the turbo code performs much better than the convolutional code, this is not the case on the block fading channel for which the latter shows a better resistance to transmission errors.

At the 4-th iteration, the turbo code over the independent Rayleigh fading channel performs 2 dB better than the convolutional code for a bit error rate 10^{-6}, as illustrated in Figure 3. Note however that the turbo code progress is less than 1 dB through iterations while the convolutional code gains more than 1.5 dB. This is due to the fact that the turbo code has its own internal decoding iterations, 4 turbo iterations for each global detection iteration.

On the block fading channel, as illustrated by Figures 4 and 5, the proposed turbo code improves by 1 dB after four global iterations but unfortunately it is 0.7 dB less efficient than the 64-state convolutional code. This moderate performance of the turbo code is mainly caused by the absence of the interleaving gain [15] when frame error rate (FER) in considered. Moreover, to perform correctly a turbo code needs a relatively

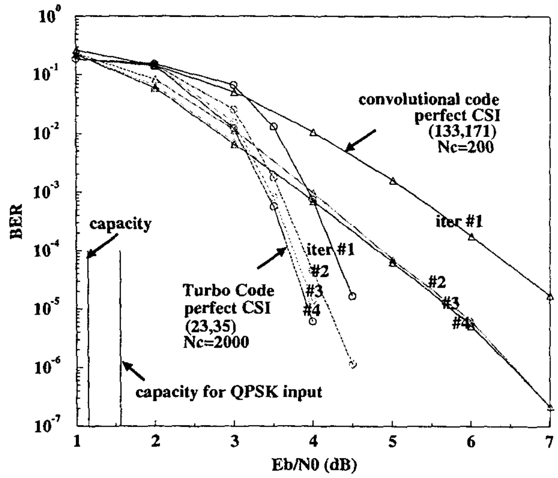

Figure 3: Bit error rate for convolutional and parallel turbo codes, independent Rayleigh fading channel, $n_t = n_r = 2$ antennas.

large frame size which makes it more sensitive to frame errors than a simple convolutional code.

All computer simulations discussed up to now assume a perfect knowledge of the channel state. The EM estimation has been applied to the convolutional code with 12 pilot symbols added to the original frame, that is to say $12/112 \approx 11\%$ pilots with $N_c = 200$ coded bits. Taking into account the spectral efficiency loss caused by the pilots insertion, the curves (with circles) shown in Figure 5 indicate a 1.5 dB loss compared to the perfect channel state information case (with triangles) for a frame error rate 10^{-3}. Notice that the simple insertion of the same percentage of pilot symbols, without the Expectation-Maximization algorithm, would have reduced the coding gain and placed the curve with pilot symbols at 2.9 dB from the perfect channel state estimation.

The iterative APP algorithm proposed in this paper has a complexity exponential with n_t, the number of transmitting antennas, which is prohibitive when n_t is greater than 4. The marginalization complexity in the APP conversion is dramatically reduced if we select the signal vector with the highest *a priori* probability and then toggle the corresponding bits one by one to generate the likelihoods. This sub-optimal APP evaluation has a linear complexity with n_t and performs as well as the exhaustive optimal evaluation.

V. Conclusions

We proposed an iterative detection and decoding scheme for systems with multiple transmitting and receiving antennas that can be applied to error-correcting codes in cases where maximum likelihood decoding is not feasible. The major interest of this scheme is that the APP derivation can be applied to any type of code, provided that a SISO decoder is available. We also showed that the estimation of the channel parameters with the EM algorithm can be integrated into the detection process. This technique may also be applied to a Rayleigh channel affected by a Doppler shift via a sliding-window SISO decoder that provides a local channel estimation.

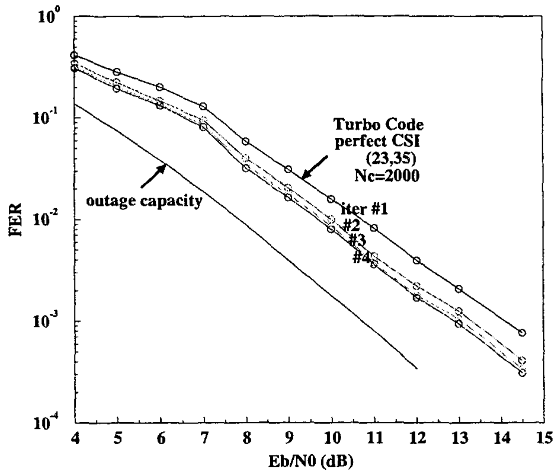

Figure 4: Frame error rate of a parallel turbo code, RSC constituent (23, 35), block fading channel, $n_t = n_r = 2$ antennas.

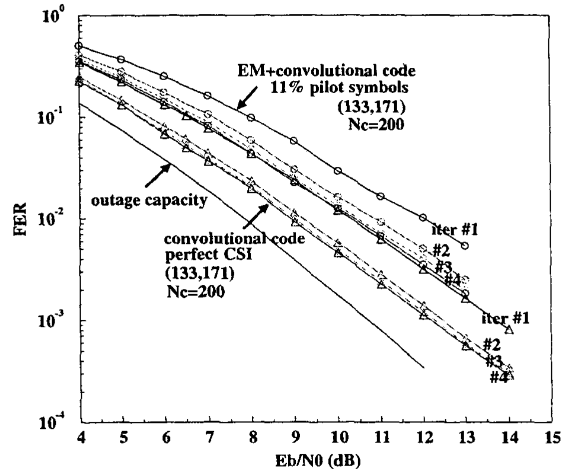

Figure 5: Frame error rate of a convolutional code, generators (133, 171), block fading channel, $n_t = n_r = 2$ antennas.

Applied to parallel turbo codes and classical convolutional codes, the iterative procedure proved to be very efficient on both Rayleigh fading and block fading channels. A small turbo code, interleaver size 1000 and total rate 1/2, exhibits a bit error rate at 2.5 dB distance from the capacity limit. A 64-state rate 1/2 convolutional code performs 1.5 dB from outage capacity with perfect channel state information and 3 dB with EM estimation on the block fading channel.

REFERENCES

[1] C. Berrou, A. Glavieux and P. Thitimajshima, "Near Shannon limit error-correcting coding and decoding: turbo-codes," *Proceedings of ICC'93*, Genve, pp. 1064–1070, May 1993.

[2] C. Berrou and A. Glavieux, "Near optimum error correcting coding and decoding: Turbo-codes," *IEEE Transactions on Communications*, vol. 44, pp. 1261–1271, October 1996.

[3] J. Hagenauer, E. Offer and L. Papke, "Iterative decoding of binary block and convolutional codes," *IEEE Transactions on Information Theory*, vol. 42, no. 2, pp. 429–445, March 1996.

[4] L. R. Bahl, J. Cocke, F. Jelinek and J. Raviv, "Optimal decoding of linear codes for minimizing symbol error rate," *IEEE Transactions on Information Theory*, vol. 20, pp. 284–287, March 1974.

[5] S. Benedetto, G. Montorsi and F Pollara, "A Soft-Input Soft-Output APP Module for Iterative Decoding of Concatenated Codes," *IEEE Communications Letters*, vol. 1, no. 1, pp. 22-24, January 1997.

[6] E. Zehavi, "8-PSK trellis codes for a rayleigh channel," *IEEE Transactions on Communications*, vol. 40, pp. 873–884, May 1992.

[7] G. Caire, G. Taricco and E. Biglieri, "Bit-interleaved coded modulation," *IEEE Transactions on Information Theory*, vol. 44, no. 3, pp.927–946, May 1998.

[8] E. Telatar, "Capacity of multi-antenna Gaussian channels," *AT&T Bell Laboratories Internal Tech. Memo.* , June 1995.

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

[10] V. Tarokh, N. Seshadri and A. R. Calderbank, "Space-time codes for high data rate wireless communication: Performance criterion and code construction," *IEEE Transactions on Information Theory*, vol. 44, no. 2, pp.744–765, March 1998.

[11] A. Stefanov and T. M. Duman, "Turbo coded modulation for systems with transmit and receive antenna diversity," *Proceedings of the 1999 Global Telecommunications Conference*, Dec. 5-9, 1999, Rio de Janeiro, Brazil, pp. 2336–2340.

[12] A. P. Dempster, N. M. Laird and D. B. Rubin, "Maximum likelihood from incomplete data via the EM algorithm," *J. Roy. Stat. Soc.* , no. 39, pp. 1–38, 1977.

[13] C. Lamy, F. Boixadera and J. Boutros, "Iterative APP decoding and channel estimation for multiple-input multiple-output channels," submitted to the *IEEE Transactions on Communications*, January 2000.

[14] S. Lin and D. J. Costello, "Error Control Coding: Fundamentals and applications," Englewood Cliffs, NJ, Prentice-Hall, 1983.

[15] S. Benedetto and G. Montorsi, "Design of parallel concatenated convolutional codes," *IEEE Transactions on Communications*, vol. 44, no. 5, pp. 591–600, May 1996.

[16] J. Boutros, "Les turbo codes parallèles et séries, décodage SISO itératif et performances ML," ENST graduate course, *http://www.com.enst.fr/turbocodes.*

[17] C. N. Georghiades and J. C. Han, "Sequence estimation in the presence of random parameters via the EM algorithm," *IEEE Transactions on Communications*, vol. 45, no. 3, pp. 300–308, March 1997.

[18] N. Ibrahim, "Codage et décodage de canal pour un système de communication à accès multiple," Ph.D. dissertation, École Nationale Suprieure des Télécommunications, Paris, France, March 1999.

[19] C. Brutel, J. Boutros and P. Mège, "Iterative joint channel estimation and detection of coded CPM," *Proceedings of the 2000 International Zürich Seminar on Broadband Communications*, Feb. 15-17 2000, ETH Zürich, Switzerland.