ITW 2002, Bangalore, India, Oct. 20-25, 2002

Application of the error impulse method in the design of high-order turbo coded modulation

Claude Berrou, Michel Jézéquel, Catherine Douillard and Laura Conde
PRACOM/ENST Bretagne, BP 832, 29285 Brest Cedex, France
e-mail: claude.berrou@enst-bretagne.fr

Abstract – **A new method for computing minimum distances, based on the ability of the soft-in decoder (if it exists) to overcome error impulse patterns, was recently proposed. This error impulse method, which delivers the results in record time, makes it possible to easily predict the asymptotic behaviour of the code considered. In this paper, we address the association of a turbo code and an *M*-PSK modulation, with the so-called pragmatic or BICM approach, and we develop a rationale to enable the performance prediction of such a coded scheme.**

I. INTRODUCTION

The growing interest for high spectral efficiency transmitting systems aiming at low residual error rates makes it compulsory to contemplate judicious associations of high-order modulations and powerful channel coding techniques. One possible solution involves combining the chosen modulation, such as 8-PSK, 16-QAM or higher order, with a turbo code, according to the so-called pragmatic approach [1], which in fact is very similar to the Bit-Interleaved Coded Modulation (BICM) concept [2], introduced more recently. Such designs face a serious problem of performance prediction at low error rates, for several reasons. First, and especially when using turbo codes, the minimum distance $d_{\min}$ of the code may not be known accurately. Even if $d_{\min}$ is available, the decoder may be not optimal, that is, the performance at low error rates may be far from the theoretical asymptote. Moreover, high-order modulations are not linear when reasoning at the bit level, and then the bounds that can be calculated when all bits are equally protected, like with BPSK or QPSK, no longer hold.

A new method for the estimation of the minimum distances of linear codes, in particular turbo codes, has recently been proposed. This method is briefly described in the next section. Section III poses the equations that lead to the estimation technique we propose on the asymptotical performance of turbo coded BICMs for a Gaussian channel, and section IV gives some examples of comparison between simulated and estimated performance.

The research reported in this paper was supported by France Télécom R&D.

II. THE ERROR IMPULSE METHOD

The minimum distance $d_{\min}$ of some error-correcting codes, in particular turbo codes and turbo-like codes, is not easy to determine by classical methods, such as the exhaustive enumeration of codewords. An original, simple and fast algorithm for the estimation of the minimum distances of linear codes, was presented in [3]. This is based on the ability of the soft-in decoder, if it is available, to overcome error impulse patterns added to the reference (all-zero) sequence. The Error Impulse Response (EIR) of a coding/decoding scheme was introduced in [4]. The encoded block is "all zero", and after the logical/antipodal conversion, only "-1" values are transmitted to the decoder, except for one symbol, to which the Error Impulse (EI) is added.

A Maximum Likelihood (ML) decoder is able to recover the emitted codeword if the scalar product it calculates from the received symbols does not deviate from the ideal one by $d_{\min}$ or more. This deviation may be imputable to one or several symbols, the latter case being the common situation in real receivers. Because the scalar product is a distributive operation when one of the vectors has n equal components, we can write:

$$\sum_{i=1,n}(-1)(-1+\delta_i) = n-1+(-1)(-1+\sum_{i=1,n}\delta_i) \qquad (1)$$

where δ_i is the deviation (noise) from the ideal "-1" value for the i_{th} received symbol ($1 \le i \le n$). Relation (1) shows that condition $\sum_{i=1,n}\delta_i < d_{\min}$ can be checked by adding an EI to any single symbol of the codeword. In practice, the method is mainly applicable to linear systematic codes and the EI, with magnitude A, may be added to any information (systematic) symbol.

Therefore, the proposed method involves applying an EI with magnitude A_i to the i_{th} information symbol, increasing A_i until the decoder fails to recover the "all-zero" codeword. We denote A_i^* the integer part of this threshold and we conclude that the minimum distance of codewords such that the i_{th} symbol is logical "1", is equal to A_i^*. The minimum distance of the code is obtained by testing all the positions i of the information bits ($1 \le i \le k$). If the code displays some periodicity, a limited number of symbols can be tested. For instance, neglecting the side effects, the number of symbols to be analysed for the turbo code is the lowest common multiple of the permutation period and the puncturing pattern period.

The principle of the EI method relies strictly on ML decoding, but experience shows that it can also be carried out successfully with sub-optimal Max-Log-MAP turbo decoding, which does not require the knowledge of the channel parameters (which is fortunate for this kind of perturbation!). The results are generally in agreement with that given by other tools, even if, in some cases (typically for short blocks), the minimum distance may be underestimated, but very rarely by more than 5%.

Nevertheless, care must be taken in making use of the method. For instance, for a classical turbo code, one has to avoid applying the EI on redundant symbols. This is because the iterative decoding procedure would be highly unbalanced, with one decoder facing the EI, and not the other.

Because the method does not explicitly provide the multiplicity of codewords of weight $d_{\min}$ and higher, we must introduce some further hypotheses to use the classical asymptotic estimate of the Frame Error Rate (FER) for an Additive White Gaussian Noise (AWGN):

$$\mathrm{FER} \approx \frac{1}{2} m(d_{\min}) \mathrm{erfc}(\sqrt{R d_{\min} \frac{E_b}{N_0}}) \tag{2}$$

where $m(d_{\min})$ is the multiplicity, that is, the number of codewords at distance $d_{\min}$, and R is the coding rate.

Hyp. 1: there is only one codeword with distance A_i^* such that the i_{th} symbol is logical "1".

Hyp. 2: all the distances A_i^* obtained for the whole set of positions i ($1 \le i \le k$) concern distinct codewords (there is no overlapping).

The former hypothesis is optimistic, while the latter is pessimistic. Taken together, as confirmed by experience, they provide a good estimate of the FER. For some codes, such as product codes for which $m(d_{\min})$ is very large, these hypotheses are irrelevant. Adopting **Hyp. 1** and **2** gives the following estimation of the asymptotic FER, which is relevant for baseband as well as for BPSK or QPSK modulated signals:

$$\mathrm{FER} \approx \frac{1}{2} \sum_{i=1,k} \mathrm{erfc}(\sqrt{R A_i^* \frac{E_b}{N_0}}) \tag{3}$$

III. HIGH ORDER CODED MODULATIONS

We consider the transmission scheme depicted in Fig. 1, based on the so-called pragmatic or BICM approach. Incoming data are first encoded by an error-correcting code which yields codewords whose bits are interleaved by permutation Π. Then a mapper puts them together in groups of m bits, following the Gray code, to form symbols suited to the modulator. On the receiver side, the projections onto the two axes enable us to calculate the Logarithms of Likelihood Ratios (LLRs) for each bit of the received block. After the inverse permutation Π^{-1}, these LLRs are used by the error-correcting decoder.

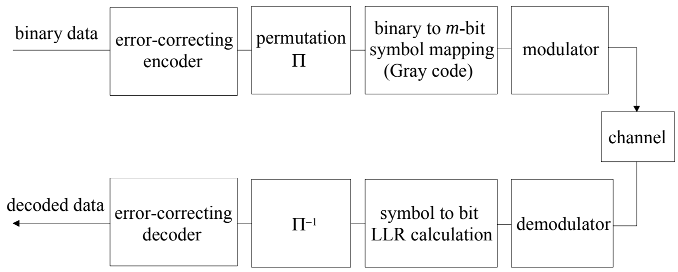

Fig. 1. Transmission scheme according to the pragmatic approach.

If the modulation chosen has more than 4 points in its constellation, it is not possible to test directly the system with an EI channel, because not all the bits are equally protected. The performance depends on the data sequence, and an "all-zero" sequence with only one erroneous value is no longer representative of the error-correcting power of the coded modulation. But the knowledge of $d_{\min}$, or at least of the estimated distances A_i^*, given by the EI method as described in the previous section, makes possible a reliable estimation of the asymptotic performance.

To be concrete, we restrict the study to M-PSK modulation ($M = 2^m$), the extension to other cases being fairly easy to imagine. The starting point of the rationale is the probability P_e for the decoder to decide in favour of a wrong codeword instead of the correct emitted codeword. One codeword is transmitted using N modulated symbols, and $\{\varphi_i\}$ and $\{\varphi_i'\}$, ($1 \le i \le N$), are the sequences of the emitted carrier phases corresponding to the correct and the wrong codewords, respectively. It can be shown for a Gaussian channel (see the annex) that P_e is equal to:

$$P_e = \frac{1}{2} \mathrm{erfc} \sqrt{\frac{E_s}{N_0} \left[\sum_{i=1,N} \sin^2\left(\frac{\varphi_i' - \varphi_i}{2} \right) \right]} \tag{4}$$

where E_s is the energy per symbol and N_0 is the one-sided noise density on the AWGN channel. If we assume the following hypothesis:

Hyp. 3: a symbol does not contain more than one opposite bit in the correct and in the wrong codewords,

the worst case for P_e is given by a minimum phase gap for each bit that differs in the codewords. Regarding only the asymptotic situation where the number of these opposite bits is $d_{\min}$, P_e is then upper-bounded by:

$$P_e \le \frac{1}{2}\operatorname{erfc}\sqrt{\frac{E_s}{N_0} d_{\min} \sin^2\left(\frac{\pi}{M}\right)} \qquad (5)$$

This gives a pessimistic estimation, because $\varphi_i' - \varphi_i$ is sometimes higher than the minimum value, the proportion and the values of the favorable cases depending on the constellation [5].

In the following, and for sake of simplicity, we will only consider the particular case of 8-PSK, whose constellation and mapping are given in Fig. 2.

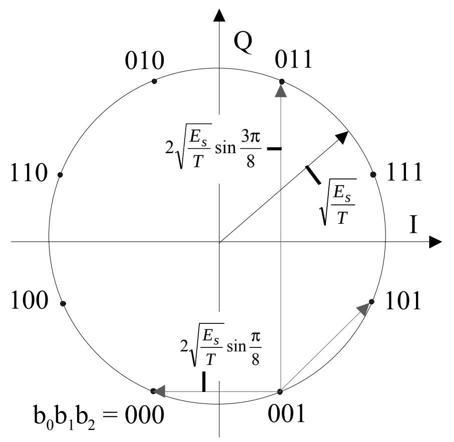

Fig. 2. 8-PSK constellation with Gray encoding. E_s and T are the symbol energy and duration.

The minimum Euclidean distance between symbols is $2\sqrt{\frac{E_s}{T}}\sin(\pi/8)$. But the Euclidean distance between one symbol and any other with only one bit different is raised to $2\sqrt{\frac{E_s}{T}}\sin(3\pi/8)$, in one case out of three (see Fig. 2). So, statistically, one third of the bits are better protected than the others. In order to benefit from this statistical property, we draw permutation Π at random, so that we can set:

$$\Pr\{\varphi_i - \varphi_i' = \pi/4\} = 2/3; \Pr\{\varphi_i - \varphi_i' = 3\pi/4\} = 1/3$$

and, using **Hyp. 3** again, relation (4) can be refined as:

$$\begin{aligned} P_{e,8-\text{PSK},\Pi\text{ random}} &= \sum_{j=0}^{d_{\min}} \binom{d_{\min}}{j} \left(\frac{1}{3}\right)^{j} \left(\frac{2}{3}\right)^{d_{\min}-j} \\ &\frac{1}{2}\operatorname{erfc}\sqrt{\frac{E_s}{N_0}\left[j\sin^2\frac{3\pi}{8} + (d_{\min}-j)\sin^2\frac{\pi}{8}\right]} \\ &= \left(\frac{2}{3}\right)^{d_{\min}} \sum_{j=0}^{d_{\min}} \binom{d_{\min}}{j} \left(\frac{1}{2}\right)^{j+1} \\ &\operatorname{erfc}\sqrt{\frac{E_s}{N_0}\left[j\sin^2\frac{3\pi}{8} + (d_{\min}-j)\sin^2\frac{\pi}{8}\right]} \end{aligned} \qquad (6)$$

For a given code with impulse spectrum $\{A_i^*\}$, and adopting **Hyp. 1** and **2**, the asymptotic FER performance evaluation is given by:

$$\begin{aligned} \text{FER}_{8\text{-PSK},\Pi\text{ random}} &\approx \sum_{i=1}^{k} \left(\frac{2}{3}\right)^{A_i^*} \sum_{j=0}^{A_i^*} \binom{A_i^*}{j} \left(\frac{1}{2}\right)^{j+1} \\ &\operatorname{erfc}\sqrt{\frac{E_s}{N_0}\left[j\sin^2\frac{3\pi}{8} + (A_i^*-j)\sin^2\frac{\pi}{8}\right]} \end{aligned} \qquad (7)$$

IV. Some Application Examples

In order to assess the relevance of (7), we consider the use of two types of turbo codes: the 8-state duo-binary turbo code adopted for the DVB-RCS/RCT standard [6,7], and its extension to 16 states [8]. For each code, four curves are plotted in Fig. 3: the upper-bound given by (5), the estimation (7) established in this paper for a BICM mapping, the corresponding simulation result, and finally the simulated performance of the pragmatic Turbo Trellis-Coded Modulation (TTCM).

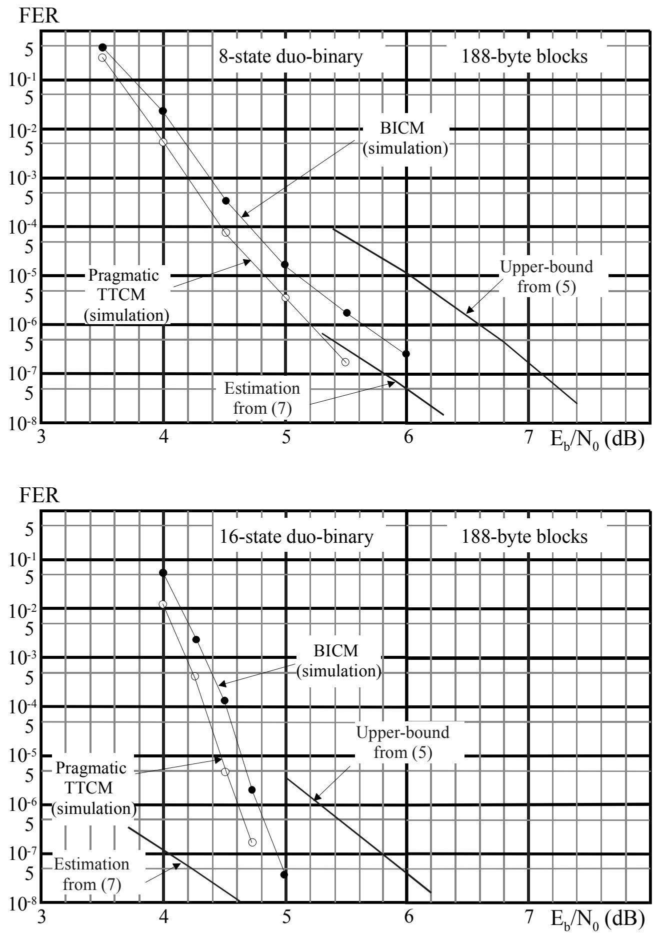

Fig. 3. Simulation results and analytical estimations, in Frame Error Rate, for 8-PSK modulation associated with 8-state and 16-state duo-binary turbo codes, on a Gaussian channel. Block length of 188 bytes (1504 bits). Decoding with the Max-Log-MAP algorithm; 8 iterations and 6-bit quantized samples. The estimation from (7) has been truncated to the lower two values of A_i^*, that is 13 (22% of bits) and 14 (65%) for the 8-state code, 18 (52%) and 19 (29%) for the 16-state code.

The pragmatic TTCM involves using the mapping proposed in [9], that is, systematically assigning the three bits stemming from each duo-binary component code to the same 8-PSK symbol, but, unlike the true TTCM, the pragmatic TTCM keeps the same turbo encoder as used in the BICM scheme, and also is based on the Gray encoding instead of the "Set Partitioning" encoding. On a Gaussian channel, the performance of TTCM, whether it is pragmatic or not, is better then that of BICM, because the minimum Euclidean distance between the correct and any wrong modulated codeword is increased. This comes from the fact that two bits, out of the three yielded by each component duo-binary encoder, are different, with an occurrence probability of ½ inside the codewords, and 1 at the extremities. However, TTCM and pragmatic TTCM perform worst than BICM on the Rayleigh channel, for which the main sought-for feature is diversity.

Regarding the comparison that can be made between the BICM simulated performance and the estimated asymptotic behaviour, according to (7), we can notice good agreement between them, in particular for the 8-state case. About 0.4 dB, which can be attributed to the iterative decoding sub-optimality [4], separates simulation and estimation curves. As for the 16-state case, it is only for very low error rates that both curves would seem to merge, apparently with a comparable residual loss of 0.4 dB.

V. CONCLUSIONS

The principle of a method for the prediction of the asymptotic performance on a Gaussian channel of pragmatic, or BICM, turbo coded modulations, has been proposed, focusing particularly on 8-PSK. This method is based on the minimum distance spectrum delivered by the error impulse technique, together with a random model of the permutation that links codewords and modulation symbols. The quality of the estimation has been assessed by a comparison with some simulated results. The study has to be pursued for other kinds of channels, mainly the Rayleigh channel, for which BICM has been proved to be advantageous.

Acknowledgement: grateful thanks to Prof. Alain Glavieux for his contribution to the writting of this paper.

ANNEX

We denote by $(X_i, Y_i), (X_i', Y_i'), (x_i, y_i)$ the quadrature components of the i-th signal ($1 \le i \le N$) in the emitted, concurrent, and received sequences, respectively. E_s and T being the energy per symbol and the symbol duration, we have:

$$X_i = \sqrt{\frac{E_S}{T}} \cos\varphi_i; \qquad Y_i = \sqrt{\frac{E_S}{T}} \sin\varphi_i$$

$$X_i' = \sqrt{\frac{E_S}{T}} \cos\varphi_i'; \qquad Y_i' = \sqrt{\frac{E_S}{T}} \sin\varphi_i'$$

We assume that reception is made by a coherent demodulator with adapted filters. Therefore,

$$x_i = X_i + b_i^I; \qquad y_i = Y_i + b_i^Q$$

where b_i^I, b_i^Q are two independent Gaussian random variables with zero mean and variance $\sigma^2 = \frac{N_0}{2T}$.

The probability of an error event when using an ML decoder (for a N-symbol sequence) is:

$$P_e = \Pr\left\{ \sum_{i=1}^{N} \left(x_i - X_i'\right)^2 + \left(y_i - Y_i'\right)^2 < \sum_{i=1}^{N} (x_i - X_i)^2 + (y_i - Y_i)^2 \right\}$$

that is

$$P_e = \Pr(Z > \lambda) = \frac{1}{\sigma_Z \sqrt{2\pi}} \int_\lambda^\infty e^{-\frac{z^2}{2\sigma_Z^2}} dz = \frac{1}{2} \operatorname{erfc}\left(\frac{\lambda}{\sigma_Z \sqrt{2}}\right)$$

where we have denoted:

$$\lambda = -\sum_{i=1}^{N} X_i (X_i' - X_i) + Y_i (Y_i' - Y_i)$$

$$Z = \sum_{i=1}^{N} b_i^I \left(X_i' - X_i\right) + b_i^Q \left(Y_i' - Y_i\right)$$

Z is a Gaussian random variable with zero mean and variance:

$$\sigma_Z^2 = \sigma^2 \sum_{i=1}^{N} (X_i' - X_i)^2 + (Y_i' - Y_i)^2$$

Finally, replacing X_i, Y_i, X_i' and Y_i' by their above values leads to relation (4).

REFERENCES

[1] S. Le Goff, A. Glavieux and C. Berrou, "Turbo-codes and high spectral efficiency modulation", *Proc. of IEEE ICC'94*, pp. 645-649, New Orleans, May 1994.
[2] G. Caire, G. Taricco, and E. Biglieri, "Bit-interleaved coded modulation", *IEEE Trans. Inform. Theory*, Vol. 44, No. 3, pp.927-946, May 1998.
[3] C. Berrou and S. Vaton, "Computing the minimum distances of linear codes by the error impulse method", to appear at *ISIT '02*, Lausanne, July 2002.
[4] C. Berrou, "Some clinical aspects of turbo codes", *1st International Symposium on Turbo Codes & Related Topics*, pp. 26-31, Brest, France, Sept. 1997.
[5] X. Li and J. M. Ritcey, "Bit-interleaved coded modulation with iterative decoding", *Proc. of ICC'99*, pp. 858-863, Vol. 2, Vancouver, June 1999.
[6] DVB, "Interaction channel for satellite distribution systems", ETSI EN 301 790, V1.2.2, pp. 21-24, Dec. 2000.
[7] DVB, "Interaction channel for digital terrestrial television", ETSI EN 301 958, V1.1.1, pp. 28-30, Aug. 2001.
[8] C. Berrou, M. Jézéquel, C. Douillard et S. Kerouédan, "The advantages of non-binary turbo codes", *ITW '2001*, Cairns, Australia, Sept. 2001.
[9] P. Robertson and T. Wörz, "Bandwidth-efficient turbo trellis-coded modulation using punctured component codes", *IEEE JSAC*, Vol. 16, N° 2, pp. 206-218, Feb. 1998.