

**NEAR SHANNON LIMIT ERROR - CORRECTING  
CODING AND DECODING : TURBO-CODES (1)**

Claude Berrou, Alain Glavieux and Punya Thitimajshima

Claude Berrou, Integrated Circuits for Telecommunication Laboratory

Alain Glavieux and Punya Thitimajshima, Digital Communication Laboratory

Ecole Nationale Supérieure des Télécommunications de Bretagne, France

(1) Patents N° 9105279 (France), N° 92460011.7 (Europe), N° 07/870,483 (USA)

**Abstract** - This paper deals with a new class of convolutional codes called *Turbo-codes*, whose performances in terms of Bit Error Rate (BER) are close to the SHANNON limit. The *Turbo-Code* encoder is built using a parallel concatenation of two Recursive Systematic Convolutional codes and the associated decoder, using a feedback decoding rule, is implemented as *P* pipelined identical elementary decoders.

**I - INTRODUCTION**

Consider a binary rate  $R=1/2$  convolutional encoder with constraint length  $K$  and memory  $M=K-1$ . The input to the encoder at time  $k$  is a bit  $d_k$  and the corresponding codeword  $C_k$  is the binary couple  $(X_k, Y_k)$  with

$$X_k = \sum_{i=0}^{K-1} g_{1i} d_{k-i} \quad \text{mod.2} \quad g_{1i} = 0,1 \quad (1a)$$

$$Y_k = \sum_{i=0}^{K-1} g_{2i} d_{k-i} \quad \text{mod.2} \quad g_{2i} = 0,1 \quad (1b)$$

where  $G_1: \{g_{1i}\}$ ,  $G_2: \{g_{2i}\}$  are the two encoder generators, generally expressed in octal form.

It is well known, that the BER of a classical Non Systematic Convolutional (NSC) code is lower than that of a classical Systematic code with the same memory  $M$  at large SNR. At low SNR, it is in general the other way round. The new class of Recursive Systematic Convolutional (RSC) codes, proposed in this paper, can be better than the best NSC code at any SNR for high code rates.

A binary rate  $R=1/2$  RSC code is obtained from a NSC code by using a feedback loop and setting one of the two outputs  $X_k$  or  $Y_k$  equal to the input bit  $d_k$ . For an RSC code, the shift register (memory) input is no longer the bit  $d_k$  but is a new binary variable  $a_k$ . If  $X_k=d_k$  (respectively  $Y_k=d_k$ ), the output  $Y_k$  (resp.  $X_k$ ) is equal to equation (1b) (resp. 1a) by substituting  $a_k$  for  $d_k$  and the variable  $a_k$  is recursively calculated as

$$a_k = d_k + \sum_{i=1}^{K-1} \gamma_i a_{k-i} \quad \text{mod.2} \quad (2)$$

where  $\gamma_i$  is respectively equal to  $g_{1i}$  if  $X_k=d_k$  and to  $g_{2i}$  if  $Y_k=d_k$ . Equation (2) can be rewritten as

$$d_k = \sum_{i=0}^{K-1} \gamma_i a_{k-i} \quad \text{mod.2.} \quad (3)$$

One RSC encoder with memory  $M=4$  obtained from an NSC encoder defined by generators  $G_1=37$ ,  $G_2=21$  is depicted in Fig.1.

Generally, we assume that the input bit  $d_k$  takes values 0 or 1 with the same probability. From equation (2), we can show that variable  $a_k$  exhibits the same statistical property

$$P_r\{a_k = 0 / a_1 = \epsilon_1, \dots, a_{k-1} = \epsilon_{k-1}\} = P_r\{d_k = \epsilon\} = 1/2 \quad (4)$$

with  $\epsilon$  is equal to

$$\epsilon = \sum_{i=1}^{K-1} \gamma_i \epsilon_i \quad \text{mod.2} \quad \epsilon = 0,1. \quad (5)$$

Thus the trellis structure is identical for the RSC code and the NSC code and these two codes have the same free distance  $d_f$ . However, the two output sequences  $\{X_k\}$  and  $\{Y_k\}$  do not correspond to the same input sequence  $\{d_k\}$  for RSC and NSC codes. This is the main difference between the two codes.

When punctured code is considered, some output bits  $X_k$  or  $Y_k$  are deleted according to a chosen puncturing pattern defined by a matrix  $P$ . For instance, starting from a rate  $R=1/2$  code, the matrix  $P$  of rate  $2/3$  punctured code is

$$P = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}$$

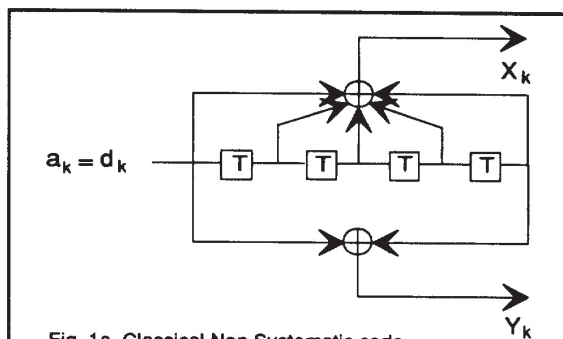


Fig. 1a Classical Non Systematic code.

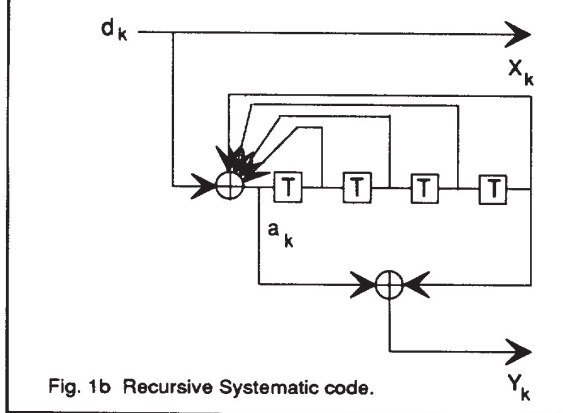
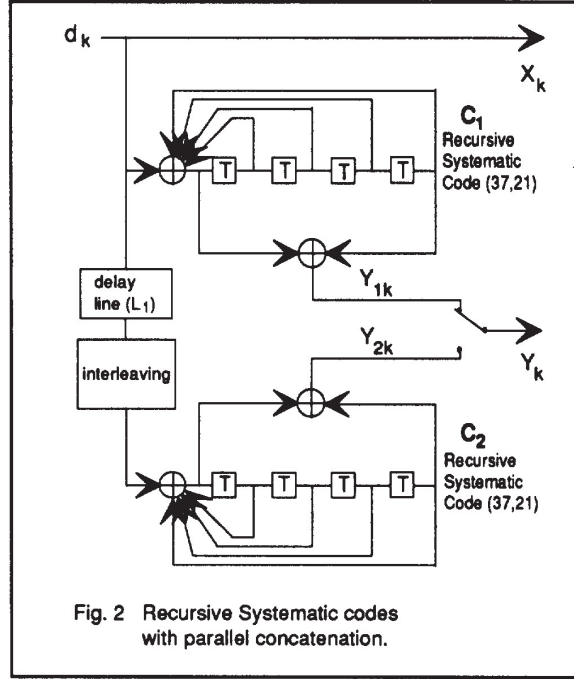


Fig. 1b Recursive Systematic code.

## II - PARALLEL CONCATENATION OF RSC CODES

With RSC codes, a new concatenation scheme, called parallel concatenation can be used. In Fig. 2, an example of two identical RSC codes with parallel concatenation is shown. Both elementary encoder ( $C_1$  and  $C_2$ ) inputs use the same bit  $d_k$  but according to a different sequence due to the presence of an interleaver. For an input bit sequence  $\{d_k\}$ , encoder outputs  $X_k$  and  $Y_k$  at time  $k$  are respectively equal to  $d_k$  (systematic encoder) and to encoder  $C_1$  output  $Y_{1k}$ , or to encoder  $C_2$  output  $Y_{2k}$ . If the coded outputs ( $Y_{1k}, Y_{2k}$ ) of encoders  $C_1$  and  $C_2$  are used respectively  $n_1$  times and  $n_2$  times and so on, the encoder  $C_1$  rate  $R_1$  and encoder  $C_2$  rate  $R_2$  are equal to

$$R_1 = \frac{n_1 + n_2}{2n_1 + n_2} \quad R_2 = \frac{n_1 + n_2}{2n_2 + n_1} \quad (6)$$



The decoder DEC depicted in Fig. 3a, is made up of two elementary decoders (DEC<sub>1</sub> and DEC<sub>2</sub>) in a serial concatenation scheme. The first elementary decoder DEC<sub>1</sub> is associated with the lower rate  $R_1$  encoder  $C_1$  and yields a soft (weighted) decision. The error bursts at the decoder DEC<sub>1</sub> output are scattered by the interleaver and the encoder delay  $L_1$  is inserted to take the decoder DEC<sub>1</sub> delay into account. Parallel concatenation is a very attractive scheme because both elementary encoder and decoder use a single frequency clock.

For a discrete memoryless gaussian channel and a binary modulation, the decoder DEC input is made up of a couple  $R_k$  of two random variables  $x_k$  and  $y_k$ , at time  $k$

$$x_k = (2d_k - 1) + i_k \quad (7a)$$

$$y_k = (2Y_k - 1) + q_k, \quad (7b)$$

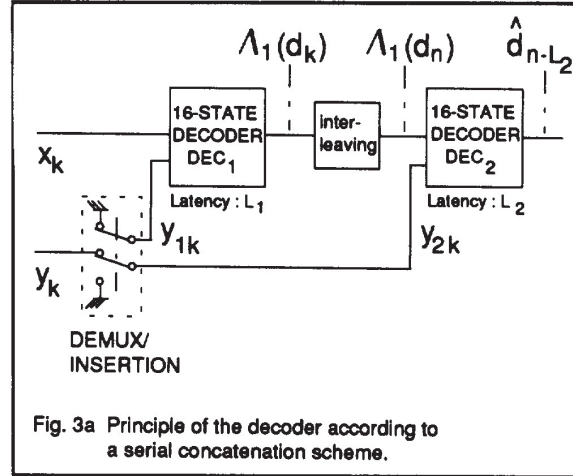
where  $i_k$  and  $q_k$  are two independent noises with the same variance  $\sigma^2$ . The redundant information  $y_k$  is demultiplexed and sent to decoder DEC<sub>1</sub> when  $Y_k = Y_{1k}$  and toward decoder DEC<sub>2</sub> when  $Y_k = Y_{2k}$ . When the redundant information of a

given encoder ( $C_1$  or  $C_2$ ) is not emitted, the corresponding decoder input is set to zero. This is performed by the DEMUX/INSERTION block.

It is well known that soft decoding is better than hard decoding, therefore the first decoder DEC<sub>1</sub> must deliver to the second decoder DEC<sub>2</sub> a weighted (soft) decision. The Logarithm of Likelihood Ratio (LLR),  $\Lambda_1(d_k)$  associated with each decoded bit  $d_k$  by the first decoder DEC<sub>1</sub> is a relevant piece of information for the second decoder DEC<sub>2</sub>

$$\Lambda_1(d_k) = \text{Log} \frac{P_r\{d_k = 1 / \text{observation}\}}{P_r\{d_k = 0 / \text{observation}\}} \quad (8)$$

where  $P_r\{d_k = i / \text{observation}\}$ ,  $i = 0, 1$  is the a posteriori probability (APP) of the data bit  $d_k$ .



## III - OPTIMAL DECODING OF RSC CODES WITH WEIGHTED DECISION

The VITERBI algorithm is an optimal decoding method which minimizes the probability of sequence error for convolutional codes. Unfortunately this algorithm is not able to yield the APP for each decoded bit. A relevant algorithm for this purpose has been proposed by BAHLE *et al.* [1]. This algorithm minimizes the bit error probability in decoding linear block and convolutional codes and yields the APP for each decoded bit. For RSC codes, the BAHLE *et al.* algorithm must be modified in order to take into account their recursive character.

### III - 1 Modified BAHLE *et al.* algorithm for RSC codes

Consider a RSC code with constraint length  $K$ ; at time  $k$  the encoder state  $S_k$  is represented by a  $K$ -uple

$$S_k = (a_k, a_{k-1}, \dots, a_{k-K+1}). \quad (9)$$

Also suppose that the information bit sequence  $\{d_k\}$  is made up of  $N$  independent bits  $d_k$ , taking values 0 and 1 with equal probability and that the encoder initial state  $S_0$  and final state  $S_N$  are both equal to zero, i.e

$$S_0 = S_N = (0, 0, \dots, 0) = 0. \quad (10)$$

The encoder output codeword sequence, noted  $C_1^N = \{C_1, \dots, C_k, \dots, C_N\}$  is the input to a discrete gaussian memoryless channel whose output is the sequence  $R_1^N = \{R_1, \dots, R_k, \dots, R_N\}$  where  $R_k = (x_k, y_k)$  is defined by relations (7a) and (7b).

The APP of a decoded data bit  $d_k$  can be derived from the joint probability  $\lambda_k^i(m)$  defined by

$$\lambda_k^i(m) = P_r \{d_k = i, S_k = m / R_1^N\} \quad (11)$$

and thus, the APP of a decoded data bit  $d_k$  is equal to

$$P_r \{d_k = i / R_1^N\} = \sum_m \lambda_k^i(m), \quad i = 0, 1. \quad (12)$$

From relations (8) and (12), the LLR  $\Lambda(d_k)$  associated with a decoded bit  $d_k$  can be written as

$$\Lambda(d_k) = \text{Log} \frac{\sum_m \lambda_k^1(m)}{\sum_m \lambda_k^0(m)}. \quad (13)$$

Finally the decoder can make a decision by comparing  $\Lambda(d_k)$  to a threshold equal to zero

$$\begin{aligned} \hat{d}_k &= 1 & \text{if } \Lambda(d_k) > 0 \\ \hat{d}_k &= 0 & \text{if } \Lambda(d_k) < 0. \end{aligned} \quad (14)$$

In order to compute the probability  $\lambda_k^i(m)$ , let us introduce the probability functions  $\alpha_k^i(m)$ ,  $\beta_k(m)$  and  $\gamma_i(R_k, m', m)$

$$\alpha_k^i(m) = \frac{P_r \{d_k = i, S_k = m, R_1^k\}}{P_r \{R_1^k\}} P_r \{d_k = i, S_k = m / R_1^k\} \quad (15)$$

$$\beta_k(m) = \frac{P_r \{R_{k+1}^N / S_k = m\}}{P_r \{R_{k+1}^N / R_1^k\}} \quad (16)$$

$$\gamma_i(R_k, m', m) = P_r \{d_k = i, R_k, S_k = m / S_{k-1} = m'\}. \quad (17)$$

The joint probability  $\lambda_k^i(m)$  can be rewritten using BAYES rule

$$\lambda_k^i(m) = \frac{P_r \{d_k = i, S_k = m, R_1^k, R_{k+1}^N\}}{P_r \{R_1^k, R_{k+1}^N\}}. \quad (18)$$

Thus we obtain

$$\lambda_k^i(m) = \frac{P_r \{d_k = i, S_k = m, R_1^k\}}{P_r \{R_1^k\}} \frac{P_r \{R_{k+1}^N / d_k = i, S_k = m, R_1^k\}}{P_r \{R_{k+1}^N / R_1^k\}}. \quad (19)$$

Taking into account that events after time  $k$  are not influenced by observation  $R_1^k$  and bit  $d_k$  if state  $S_k$  is known, the probability  $\lambda_k^i(m)$  is equal

$$\lambda_k^i(m) = \alpha_k^i(m) \beta_k(m). \quad (20)$$

The probabilities  $\alpha_k^i(m)$  and  $\beta_k(m)$  can be recursively calculated from probability  $\gamma_i(R_k, m', m)$ . From annex I, we obtain

$$\alpha_k^i(m) = \frac{\sum_{m' j=0}^1 \gamma_i(R_k, m', m) \alpha_{k-1}^j(m')}{\sum_m \sum_{m' i=0}^1 \sum_{j=0}^1 \gamma_i(R_k, m', m) \alpha_{k-1}^j(m')} \quad (21)$$

and

$$\beta_k(m) = \frac{\sum_{m' i=0}^1 \gamma_i(R_{k+1}, m, m') \beta_{k+1}(m')}{\sum_m \sum_{m' i=0}^1 \sum_{j=0}^1 \gamma_i(R_{k+1}, m', m) \alpha_k^j(m')}. \quad (22)$$

The probability  $\gamma_i(R_k, m', m)$  can be determined from transition probabilities of the discrete gaussian memoryless

channel and transition probabilities of the encoder trellis. From relation (17),  $\gamma_i(R_k, m', m)$  is given by

$$\begin{aligned} \gamma_i(R_k, m', m) &= p(R_k / d_k = i, S_k = m, S_{k-1} = m') \\ &= q(d_k = i / S_k = m, S_{k-1} = m') \pi(S_k = m / S_{k-1} = m') \end{aligned} \quad (23)$$

where  $p(\cdot)$  is the transition probability of the discrete gaussian memoryless channel. Conditionally to  $(d_k = i, S_k = m, S_{k-1} = m')$ ,  $x_k$  and  $y_k$  are two uncorrelated gaussian variables and thus we obtain

$$p(R_k / d_k = i, S_k = m, S_{k-1} = m') =$$

$$p(x_k / d_k = i, S_k = m, S_{k-1} = m')$$

$$p(y_k / d_k = i, S_k = m, S_{k-1} = m'). \quad (24)$$

Since the convolutional encoder is a deterministic machine,  $q(d_k = i / S_k = m, S_{k-1} = m')$  is equal to 0 or 1. The transition state probabilities  $\pi(S_k = m / S_{k-1} = m')$  of the trellis are defined by the encoder input statistic. Generally,  $P_r \{d_k = 1\} = P_r \{d_k = 0\} = 1/2$  and since there are two possible transitions from each state,  $\pi(S_k = m / S_{k-1} = m') = 1/2$  for each of these transitions.

#### Different steps of modified BAML *et al.* algorithm

-Step 0 : Probabilities  $\alpha_0^i(m)$  and  $\beta_N(m)$  are initialized according to relation (12)

$$\alpha_0^i(0) = 1 \quad \alpha_0^i(m) = 0 \quad \forall m \neq 0, \quad i = 0, 1 \quad (25a)$$

$$\beta_N(0) = 1 \quad \beta_N(m) = 0 \quad \forall m \neq 0. \quad (25b)$$

-Step 1 : For each observation  $R_k$ , the probabilities  $\alpha_k^i(m)$  and  $\gamma_i(R_k, m', m)$  are computed using relations (21) and (23) respectively.

-Step 2 : When the sequence  $R_1^N$  has been completely received, probabilities  $\beta_k(m)$  are computed using relation (22), and probabilities  $\alpha_k^i(m)$  and  $\beta_k(m)$  are multiplied in order to obtain  $\lambda_k^i(m)$ . Finally the LLR associated with each decoded bit  $d_k$  is computed from relation (13).

#### IV- THE EXTRINSIC INFORMATION OF THE RSC DECODER

In this chapter, we will show that the LLR  $\Lambda(d_k)$  associated with each decoded bit  $d_k$ , is the sum of the LLR of  $d_k$  at the decoder input and of another information called extrinsic information, generated by the decoder.

Using the LLR  $\Lambda(d_k)$  definition (13) and relations (20) and (21), we obtain

$$\Lambda(d_k) = \text{Log} \frac{\sum_m \sum_{m' j=0}^1 \gamma_1(R_k, m', m) \alpha_{k-1}^j(m') \beta_k(m)}{\sum_m \sum_{m' j=0}^1 \gamma_0(R_k, m', m) \alpha_{k-1}^j(m') \beta_k(m)}. \quad (26)$$

Since the encoder is systematic ( $X_k = d_k$ ), the transition probability  $p(x_k / d_k = i, S_k = m, S_{k-1} = m')$  in expression  $\gamma_i(R_k, m', m)$  is independent of state values  $S_k$  and  $S_{k-1}$ . Therefore we can factorize this transition probability in the numerator and in the denominator of relation (26)

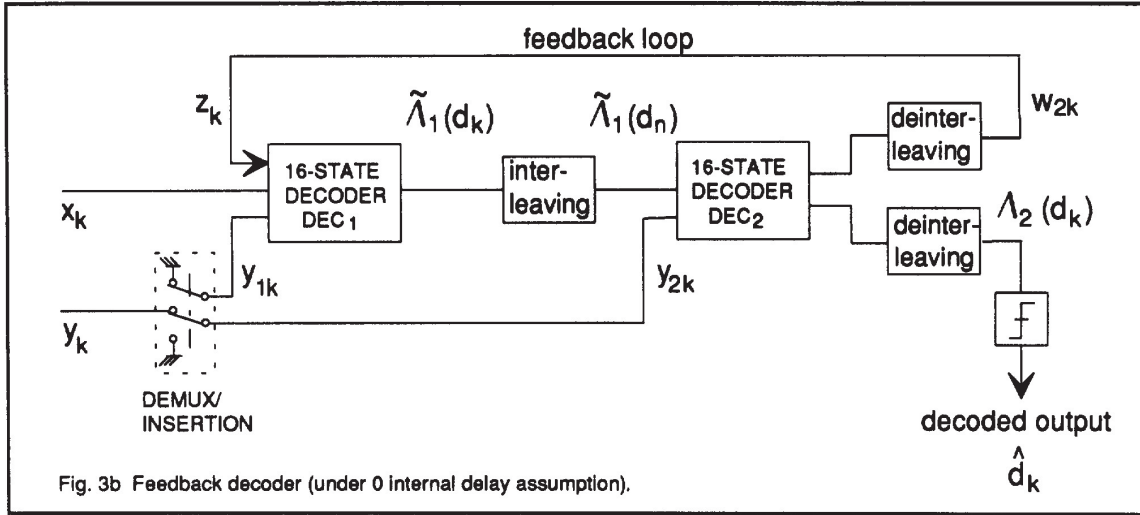


Fig. 3b Feedback decoder (under 0 internal delay assumption).

$$\Lambda(d_k) = \text{Log} \frac{p(x_k/d_k=1)}{p(x_k/d_k=0)} + \frac{\sum_{m,m',j=0}^1 \gamma_1(y_k, m', m) \alpha_{k-1}^j(m') \beta_k(m)}{\sum_{m,m',j=0}^1 \gamma_0(y_k, m', m) \alpha_{k-1}^j(m') \beta_k(m)} \quad (27)$$

Conditionally to  $d_k=1$  (resp.  $d_k=0$ ), variables  $x_k$  are gaussian with mean 1 (resp. -1) and variance  $\sigma^2$ , thus the LLR  $\Lambda(d_k)$  is still equal to

$$\Lambda(d_k) = \frac{2}{\sigma^2} x_k + W_k \quad (28)$$

where

$$W_k = \Lambda(d_k) \Big|_{x_k=0} = \frac{\sum_{m,m',j=0}^1 \gamma_1(y_k, m', m) \alpha_{k-1}^j(m') \beta_k(m)}{\sum_{m,m',j=0}^1 \gamma_0(y_k, m', m) \alpha_{k-1}^j(m') \beta_k(m)} \quad (29)$$

$W_k$  is a function of the redundant information introduced by the encoder. In general  $W_k$  has the same sign as  $d_k$ ; therefore  $W_k$  may improve the LLR associated with each decoded data bit  $d_k$ . This quantity represents the extrinsic information supplied by the decoder and does not depend on decoder input  $x_k$ . This property will be used for decoding the two parallel concatenated encoders.

## V - DECODING SCHEME OF PARALLEL CONCATENATION CODES

In the decoding scheme represented in Fig. 3a, decoder DEC<sub>1</sub> computes LLR  $\Lambda_1(d_k)$  for each transmitted bit  $d_k$  from sequences  $\{x_k\}$  and  $\{y_k\}$ , then the decoder DEC<sub>2</sub> performs the decoding of sequence  $\{d_k\}$  from sequences  $\{\Lambda_1(d_k)\}$  and  $\{y_k\}$ . Decoder DEC<sub>1</sub> uses the modified Bahl *et al.* algorithm and decoder DEC<sub>2</sub> may use the VITERBI algorithm. The global decoding rule is not optimal because the first decoder uses only a fraction of the available redundant information. Therefore it is possible to improve the performance of this serial decoder by using a feedback loop.

### V-1 Decoding with a feedback loop

We consider now that both decoders DEC<sub>1</sub> and DEC<sub>2</sub> use the modified Bahl *et al.* algorithm. We have seen in section IV that the LLR at the decoder output can be expressed as a sum of two terms if the decoder inputs were independent. Hence if the decoder DEC<sub>2</sub> inputs  $\Lambda_1(d_k)$  and  $y_{2k}$  are independent, the LLR  $\Lambda_2(d_k)$  at the decoder DEC<sub>2</sub> output can be written as

$$\Lambda_2(d_k) = f(\Lambda_1(d_k)) + W_{2k} \quad (30)$$

with

$$\Lambda_1(d_k) = \frac{2}{\sigma^2} x_k + W_{1k} \quad (31)$$

From relation (29), we can see that the decoder DEC<sub>2</sub> extrinsic information  $W_{2k}$  is a function of the sequence  $\{\Lambda_1(d_n)\}_{n \neq k}$ . Since  $\Lambda_1(d_n)$  depends on observation  $R_1^N$ , extrinsic information  $W_{2k}$  is correlated with observations  $x_k$  and  $y_{1k}$ . Nevertheless from relation (29), the greater  $|n-k|$  is, the less correlated are  $\Lambda_1(d_n)$  and observations  $x_k, y_k$ . Thus, due to the presence of interleaving between decoders DEC<sub>1</sub> and DEC<sub>2</sub>, extrinsic information  $W_{2k}$  and observations  $x_k, y_{1k}$  are weakly correlated. Therefore extrinsic information  $W_{2k}$  and observations  $x_k, y_{1k}$  can be jointly used for carrying out a new decoding of bit  $d_k$ , the extrinsic information  $z_k = W_{2k}$  acting as a diversity effect in an iterative process.

In Fig. 3b, we have depicted a new decoding scheme using the extrinsic information  $W_{2k}$  generated by decoder DEC<sub>2</sub> in a feedback loop. This decoder does not take into account the different delays introduced by decoder DEC<sub>1</sub> and DEC<sub>2</sub> and a more realistic decoding structure will be presented later.

The first decoder DEC<sub>1</sub> now has three data inputs,  $(x_k, y_{1k}, z_k)$  and probabilities  $\alpha_{1k}^j(m)$  and  $\beta_{1k}(m)$  are computed in substituting  $R_k = \{x_k, y_{1k}\}$  by  $R_k = \{x_k, y_{1k}, z_k\}$  in relations (21) and (22). Taking into account that  $z_k$  is weakly correlated with  $x_k$  and  $y_{1k}$  and supposing that  $z_k$  can be approximated by a gaussian variable with variance  $\sigma_z^2 \neq \sigma^2$ , the transition probability of the discrete gaussian memoryless channel can be now factored in three terms

$$p(R_k / d_k = i, S_k = m, S_{k-1} = m') = p(x_k / \cdot) p(y_k / \cdot) p(z_k / \cdot) \quad (32)$$

The encoder  $C_1$  with initial rate  $R_1$ , through the feedback loop, is now equivalent to a rate  $R'_1$  encoder with

$$R'_1 = \frac{R_1}{1 + R_1} \quad (33)$$

The first decoder obtains an additional redundant information with  $z_k$  that may significantly improve its performances; the term *Turbo-codes* is given for this iterative decoder scheme with reference to the turbo engine principle.

With the feedback decoder, the LLR  $\Lambda_1(d_k)$  generated by decoder  $DEC_1$  is now equal to

$$\Lambda_1(d_k) = \frac{2}{\sigma^2} x_k + \frac{2}{\sigma_z^2} z_k + W_{1k} \quad (34)$$

where  $W_{1k}$  depends on sequence  $\{z_n\}_{n \neq k}$ . As indicated above, information  $z_k$  has been built by decoder  $DEC_2$  at the previous decoding step. Therefore  $z_k$  must not be used as input information for decoder  $DEC_2$ . Thus decoder  $DEC_2$  input sequences at step  $p$  ( $p \geq 2$ ) will be sequences  $\{\tilde{\Lambda}_1(d_n)\}$  and  $\{y_{2k}\}$  with

$$\tilde{\Lambda}_1(d_n) = \Lambda_1(d_n)_{z_n=0} \quad (35)$$

Finally from relation (30), decoder  $DEC_2$  extrinsic information  $z_k = W_{2k}$ , after deinterleaving, can be written as

$$z_k = W_{2k} = \Lambda_2(d_k)_{\tilde{\Lambda}_1(d_k)=0} \quad (36)$$

and the decision at the decoder  $DEC$  output is

$$\hat{d}_k = \text{sign}[\Lambda_2(d_k)] \quad (37)$$

The decoding delays introduced by decoder  $DEC$  ( $DEC=DEC_1+DEC_2$ ), the interleaver and the deinterleaver imply that the feedback information  $z_k$  must be used through an iterative process as represented in Fig. 4a, 4b. In fact, the global decoder circuit is composed of  $P$  pipelined identical elementary decoders (Fig. 4a). The  $p$ th decoder  $DEC$  (Fig. 4b) input, is made up of demodulator output sequences  $(x)_p$  and  $(y)_p$  through a delay line and of extrinsic information  $(z)_p$  generated by the  $(p-1)$ th decoder  $DEC$ . Note that the variance  $\sigma_z^2$  of the extrinsic information and the variance of  $\tilde{\Lambda}_1(d_k)$  must be estimated at each decoding step  $p$ .

### V-2 Interleaving

The interleaver uses a square matrix and bits  $\{d_k\}$  are written row by row and read pseudo-randomly. This non-uniform reading rule is able to spread the residual error blocks of rectangular form, that may set up in the interleaver located behind the first decoder  $DEC_1$ , and to give the greater free distance as possible to the concatenated (parallel) code.

## VI - RESULTS

For a rate  $R=1/2$  encoder with constraint length  $K=5$ , generators  $G_1=37$ ,  $G_2=21$  and parallel concatenation ( $R_1=R_2=2/3$ ), we have computed the Bit Error Rate (BER) after each decoding step using the Monte Carlo method, as a function of signal to noise ratio  $E_b/N_0$  where  $E_b$  is the energy received per information bit  $d_k$  and  $N_0$  is the noise monolateral power spectral density. The interleaver consists of a  $256 \times 256$  matrix and the modified BAHAL *et al.* algorithm has been used with length data block of  $N=65536$  bits. In

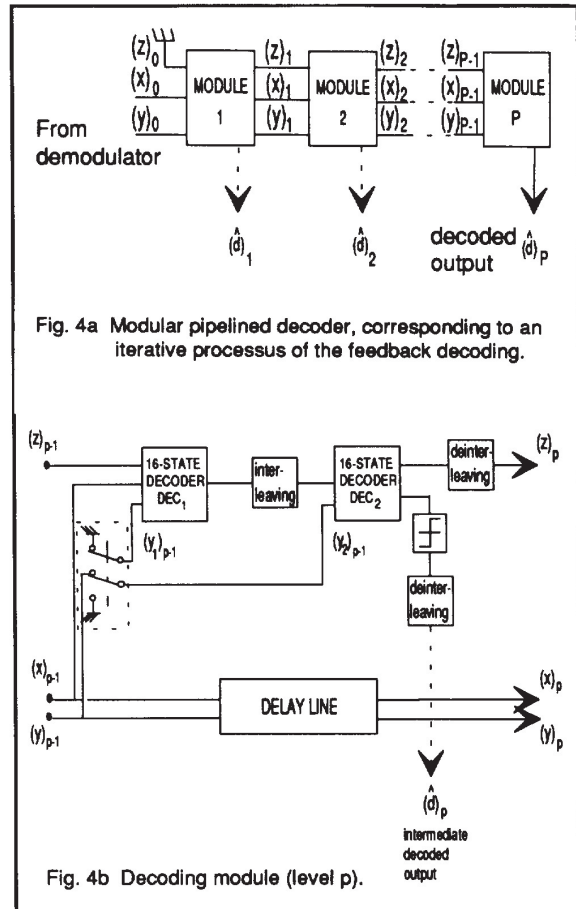


Fig. 4a Modular pipelined decoder, corresponding to an iterative process of the feedback decoding.

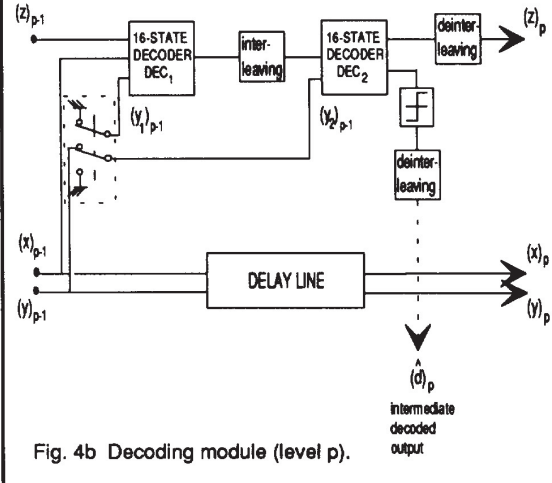
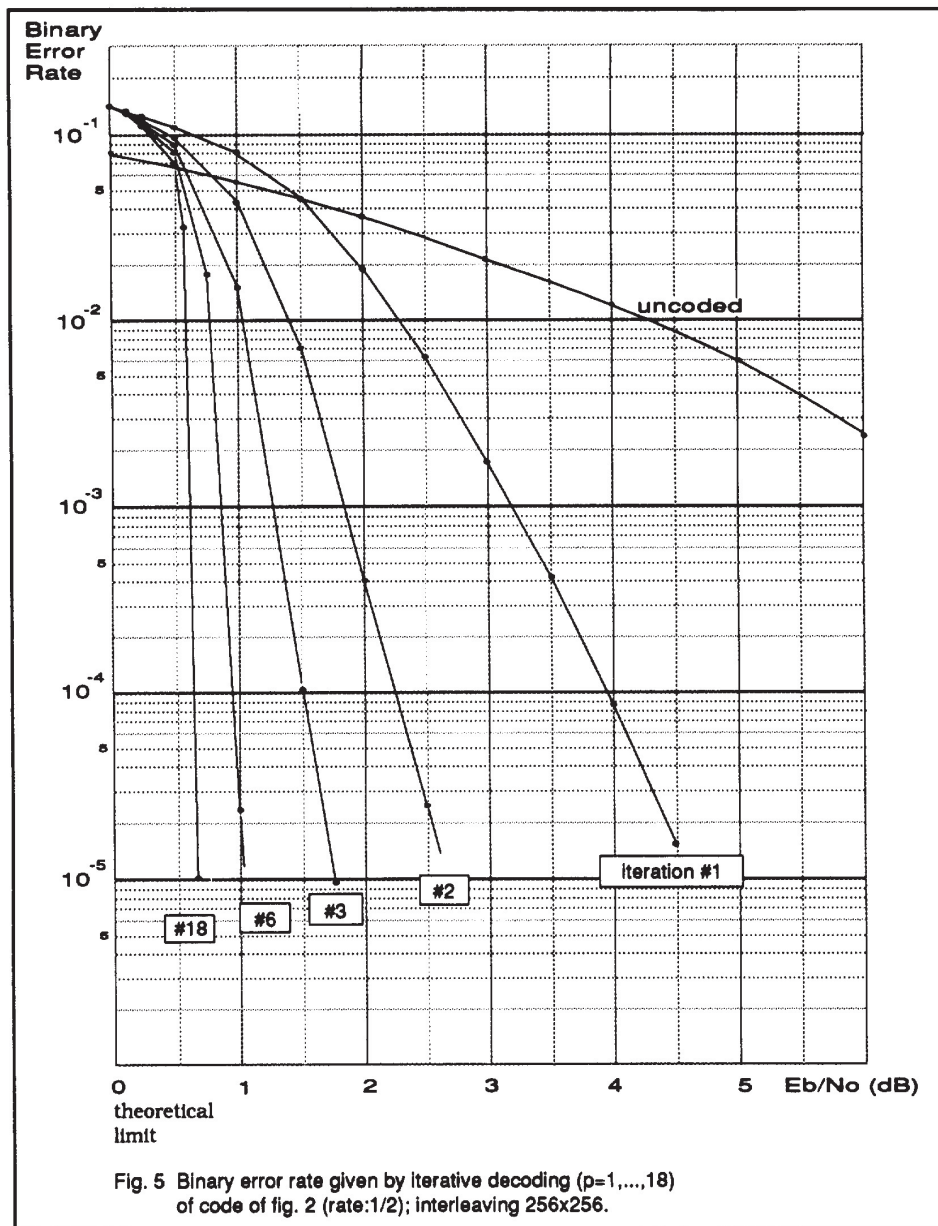


Fig. 4b Decoding module (level p).

order to evaluate a BER equal to  $10^{-5}$ , we have considered 128 data blocks i.e. approximately  $8 \times 10^6$  bits  $d_k$ . The BER versus  $E_b/N_0$ , for different values of  $p$  is plotted in Fig. 5. For any given signal to noise ratio greater than 0 dB, the BER decreases as a function of the decoding step  $p$ . The coding gain is fairly high for the first values of  $p$  ( $p=1,2,3$ ) and carries on increasing for the subsequent values of  $p$ . For  $p=18$  for instance, the BER is lower than  $10^{-5}$  at  $E_b/N_0=0,7$  dB. Remember that the Shannon limit for a binary modulation with  $R=1/2$ , is  $P_e = 0$  (several authors take  $P_e=10^{-5}$  as a reference) for  $E_b/N_0=0$  dB. With parallel concatenation of RSC convolutional codes and feedback decoding, the performances are at 0,7 dB from Shannon's limit.

The influence of the constraint length on the BER has also been examined. For  $K$  greater than 5, at  $E_b/N_0=0,7$  dB, the BER is slightly worst at the first ( $p=1$ ) decoding step and the feedback decoding is inefficient to improve the final BER. For  $K$  smaller than 5, at  $E_b/N_0=0,7$  dB, the BER is slightly better at the first decoding step than for  $K$  equal to 5, but the correction capacity of encoders  $C_1$  and  $C_2$  is too weak to improve the BER with feedback decoding. For  $K=4$  (i.e. 8-state elementary decoders) and after iteration 18, a BER of  $10^{-5}$  is achieved at  $E_b/N_0=0,9$  dB. For  $K$  equal to 5, we have tested several generators ( $G_1, G_2$ ) and the best results were achieved with  $G_1=37, G_2=21$ .



For low signal to noise ratios, we have sometimes noticed that BER could increase during the iterative decoding process. In order to overcome this effect, we have divided the extrinsic information  $z_k$  by  $[1 + \theta |\tilde{\Lambda}_1(d_k)|]$  with  $\theta = 0,15$ .

In Fig. 6, the histogram of extrinsic information  $(z)_p$  has been drawn for several values of iteration  $p$ , with all data bits equal to 1 and for a low signal to noise ratio ( $E_b/N_0 = 0,8$  dB). For  $p=1$  (first iteration), extrinsic information  $(z)_p$  is very poor about bit  $d_k$ , furthermore the gaussian hypothesis made above for extrinsic information  $(z)_p$ , is not satisfied! Nevertheless when iteration  $p$  increases, the histogram merges towards a gaussian law with a mean equal to 1. For instance, for  $p=13$ , extrinsic information  $(z)_p$  becomes relevant information concerning data bits.

## VII CONCLUSION

In this paper, we have presented a new class of convolutional codes called *Turbo-codes* whose performances in terms of BER are very close to SHANNON's limit. The decoder is made up of  $P$  pipelined identical elementary modules and rank  $p$  elementary module uses the data information coming from the demodulator and the extrinsic information generated by the rank  $(p-1)$  module. Each elementary module uses a modified Bahl *et al.* algorithm which is rather complex. A much simpler algorithm yielding weighted (soft) decisions has also been investigated for *Turbo-codes* decoding [2], whose complexity is only twice the complexity of the VITERBI algorithm, and with performances which are very close to those of the Bahl *et al.* algorithm. This new algorithm will enable encoders and

decoders to be integrated in silicon with error correcting performances unmatched at the present time.

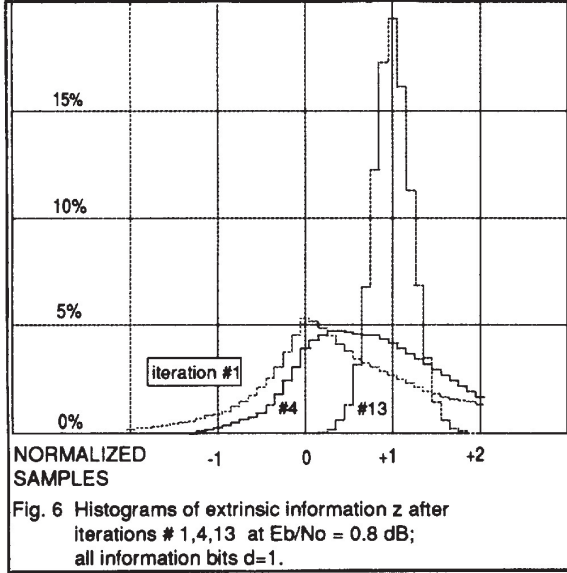


Fig. 6 Histograms of extrinsic information  $z$  after iterations #1,4,13 at  $E_b/N_0 = 0.8$  dB; all information bits  $d=1$ .

ANNEX I : EVALUATION OF PROBABILITIES  $\alpha_k^i(m)$  AND  $\beta_k(m)$ .

From relation (15) probability  $\alpha_k^i(m)$  is equal to

$$\alpha_k^i(m) = \frac{Pr\{d_k = i, S_k = m, R_1^{k-1}, R_k\}}{Pr\{R_1^{k-1}, R_k\}} = \frac{Pr\{d_k = i, S_k = m, R_k / R_1^{k-1}\}}{Pr\{R_k / R_1^{k-1}\}}. \quad (A1)$$

The numerator of  $\alpha_k^i(m)$  can be expressed from state  $S_{k-1}$  and bit  $d_{k-1}$ .

$$Pr\{d_k = i, S_k = m, R_k / R_1^{k-1}\} = \sum_{m'} \sum_{j=0}^1 Pr\{d_k = i, d_{k-1} = j, S_k = m, S_{k-1} = m', R_k / R_1^{k-1}\} \quad (A2)$$

By using BAYES rule, we can write

$$Pr\{d_k = i, S_k = m, R_k / R_1^{k-1}\} = \sum_{m'} \sum_{j=0}^1 \frac{Pr\{d_{k-1} = j, S_{k-1} = m', R_1^{k-1}\}}{Pr\{R_1^{k-1}\}}$$

$$Pr\{d_k = i, S_k = m, R_k / R_1^{k-1}\} = Pr\{d_{k-1} = j, S_{k-1} = m', R_1^{k-1}\}. \quad (A3)$$

By taking into account that events after time  $(k-1)$  are not influenced by observation  $R_1^{k-1}$  and bit  $d_{k-1}$  if state  $S_{k-1}$  is known and from relation (17) we obtain

$$Pr\{d_k = i, S_k = m, R_k / R_1^{k-1}\} = \sum_{m'} \sum_{j=0}^1 \gamma_i(R_k, m' m) \alpha_{k-1}^j(m'). \quad (A4)$$

The denominator can be also expressed from bit  $d_k$  and state  $S_k$

$$Pr\{R_k / R_1^{k-1}\} = \sum_{m'} \sum_{i=0}^1 Pr\{d_k = i, S_k = m, R_k / R_1^{k-1}\} \quad (A5)$$

and from relation (A4), we can write :

$$Pr\{R_k / R_1^{k-1}\} = \sum_m \sum_{m'} \sum_{i=0}^1 \sum_{j=0}^1 \gamma_i(R_k, m' m) \alpha_{k-1}^j(m'). \quad (A6)$$

Finally probability  $\alpha_k^i(m)$  can be expressed from probability  $\alpha_{k-1}^j(m')$  by the following relation

$$\alpha_k^i(m) = \frac{\sum_{m'} \sum_{j=0}^1 \gamma_i(R_k, m' m) \alpha_{k-1}^j(m')}{\sum_m \sum_{m'} \sum_{i=0}^1 \sum_{j=0}^1 \gamma_i(R_k, m' m) \alpha_{k-1}^j(m')}. \quad (A7)$$

In the same way, probability  $\beta_k(m)$  can be recursively calculated from probability  $\beta_{k+1}(m)$ . From relation (16), we have

$$\beta_k(m) = \frac{Pr\{R_{k+1}^N / S_k = m\}}{Pr\{R_{k+1}^N / R_1^k\}} = \frac{\sum_{m', i=0}^1 \sum_{j=0}^1 Pr\{d_{k+1} = i, S_{k+1} = m', R_{k+2}^N, R_{k+1} / S_k = m\}}{Pr\{R_{k+1}^N / R_1^k\}}. \quad (A8)$$

By using BAYES rule, the numerator is equal to

$$Pr\{R_{k+1}^N / S_k = m\} = \sum_{m'} \sum_{i=0}^1 Pr\{R_{k+2}^N / S_{k+1} = m'\} Pr\{d_{k+1} = i, S_{k+1} = m', R_{k+1} / S_k = m\}. \quad (A9)$$

By taking into account expressions of  $\gamma_i(R_{k+1}, m, m')$  and  $\beta_{k+1}(m')$ , we can write

$$\beta_k(m) = \frac{\sum_{m'} \sum_{i=0}^1 \gamma_i(R_{k+1}, m, m') \beta_{k+1}(m')}{Pr\{R_{k+1} / R_1^k\}}. \quad (A10)$$

In substituting  $k$  by  $(k+1)$  in relation (A6), the denominator of (A10) is equal to

$$Pr\{R_{k+1} / R_1^k\} = \sum_m \sum_{m'} \sum_{i=0}^1 \sum_{j=0}^1 \gamma_i(R_{k+1}, m' m) \alpha_k^j(m'). \quad (A11)$$

Finally probability  $\beta_k(m)$  can be expressed from probability  $\beta_{k+1}(m')$ , by the following relation

$$\beta_k(m) = \frac{\sum_{m'} \sum_{i=0}^1 \gamma_i(R_{k+1}, m, m') \beta_{k+1}(m')}{\sum_m \sum_{m'} \sum_{i=0}^1 \sum_{j=0}^1 \gamma_i(R_{k+1}, m' m) \alpha_k^j(m')}. \quad (A12)$$

REFERENCES

- [1] L.R. Bahl, J. Cocke, F. Jeinek and J. Raviv, "Optimal decoding of linear codes for minimizing symbol error rate", *IEEE Trans. Inform. Theory*, vol. IT-20, pp. 248-287, March 1974.
- [2] C. Berrou, P. Adde, E. Angui and S. Faudeil, "A low complexity soft-output Viterbi decoder architecture", to appear at *ICC '93*.