# Chapter 10

# Turbo codes and large spectral efficiency transmissions

Transporting information in telecommunication systems is carried out at higher and higher data rates and in narrower and narrower frequency bands. Consequently, we wish to maximize the ratio of the useful data rate to bandwidth, that is to say, the spectral efficiency of the transmissions. To do this, it seems natural to couple digital modulations having large constellations with powerful high-rate error correcting codes like turbo codes.

The studies undertaken in this domain are essentially based on two approaches: *turbo trellis coded modulation* and *pragmatic turbo coded modulation*.

## 10.1 Turbo trellis coded modulation (TTCM)

Turbo trellis coded modulation or TTCM was introduced by Robertson and Wörz in 1995 [10.5, 10.6]. It uses the notion of parallel concatenation, which is at the origin of turbo coding, applied to two trellis coded modulations, or TCMs.

TCM, introduced by Ungerboeck at the beginning of the 80s [10.7] is based on the joint optimization of error correction coding and modulation. The coding is performed directly in the signal space so that the error correcting code and the bit to signal mapping of the modulation can be represented jointly using a single trellis. The criterion for optimizing a TCM thus involves maximizing the minimum Euclidean distance between two coded sequences. To do this, Ungerboeck proposed a two-step approach: partitioning the constellation of the modulation into sub-constellations presenting increasing minimum Euclidean distances, then assigning to each branch of the trellis a signal belonging to the constellation, respecting a set of rules such as those described in [10.7].

The TTCM scheme presented by Robertson and Wörz is shown in Figure 10.1. Each TCM encoder is made up of a recursive systematic convolutional encoder, or RSC encoder, with rate $q/(q+1)$, and a modulation without memory of order $Q = 2^{q+1}$. The binary symbols coming from the source are grouped into symbols of $q$ bits. These symbols are encoded by the first TCM in the order in which are produced by the source and by the second TCM after interleaving.

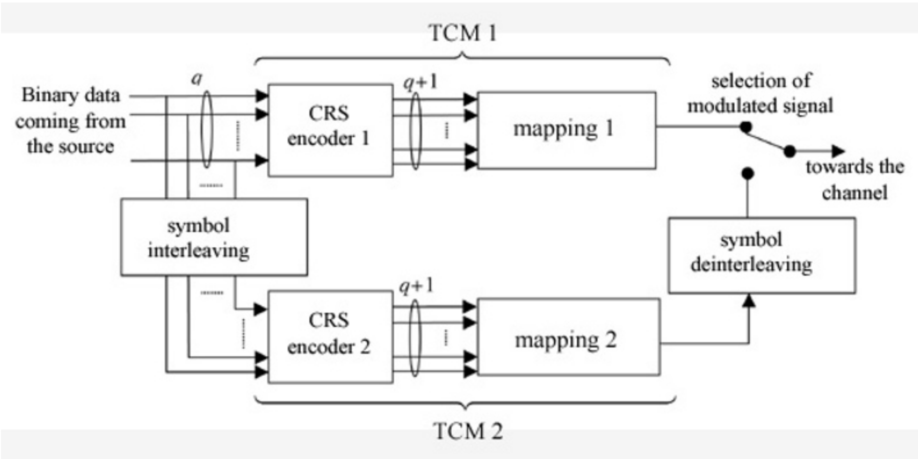



Figure 10.1 – Diagram of the principle of turbo trellis coded modulation (TTCM), according to Robertson and Wörz [10.5, 10.6]. Spectral efficiency $\eta = q$ bit/s/Hz.

Each $q$-tuple coming from the source being encoded two times, a selection operator alternatively transmits the output of one of the two TCM encoders, in order to avoid the double transmission of information, which would lead to a spectral efficiency of the system of $q/2$ bit/s/Hz. This in fact amounts to puncturing half of the redundancy sequence for each convolutional code.

At reception, the TTCM decoder is similar to a turbo decoder, except that the former directly processes the $(q + 1)$-ary symbols coming from the demodulator. Thus, the calculation of the transition probabilities at each step of the MAP algorithm (see Section 7.4) uses the Euclidean distance between the received symbol and the symbol carried by each branch of the trellis. If the decoding algorithm operates in the logarithmic domain (Log-MAP, Max-Log-MAP), it is the branch metrics that are taken equal to the Euclidean distances. Computing an estimate of the bits carried by each demodulated symbol, before decoding, would indeed be a sub-optimal implementation of the receiver.

Similarly, for efficient implementation of turbo decoding, the extrinsic information exchanged by the elementary decoders must directly concern the $q$-tuples of information transmitted and not the binary elements that they are made up of. At each decoding instant, the elementary decoders thus exchange $2^q$ values of extrinsic information.

Figure 10.2 provides two examples of elementary RSC codes used in [10.5, 10.6] to build an 8-PSK TTCM with 8 states of spectral efficiency $\eta = 2$ bit/s/Hz and a 16-QAM TTCM with spectral efficiency $\eta = 3$ bit/s/Hz.

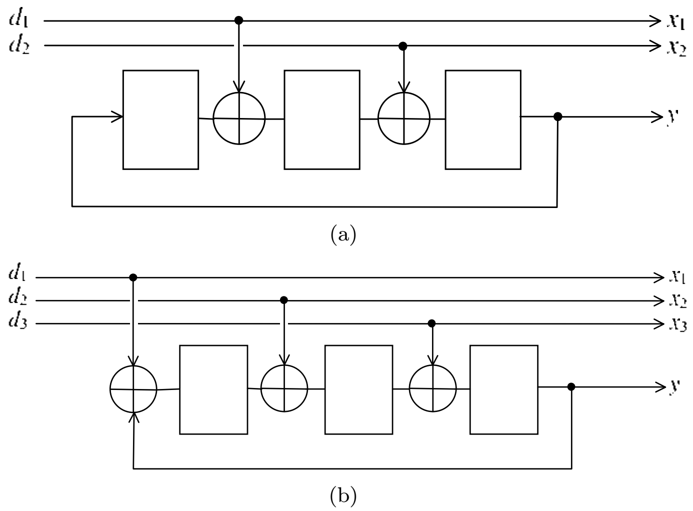



Figure 10.2 – Examples of elementary RSC codes used in [10.5, 10.6] for the construction of a 8-PSK turbo trellis (a) and a 16-QAM turbo trellis (b) coded modulations.

Figures 10.3 and 10.4 show the performance of these two TTCMs in terms of binary error rates (BER) as a function of the signal to noise ratio for transmission over a Gaussian channel. At high and average error rates, these schemes show correction performance close to capacity: a BER of $10^{-4}$ is reached at around 0.65 dB from Shannon's theoretical limit for the transmission of packets of 5,000 coded modulated symbols. On the other hand, as the interleaving function of the TTCM has not been the object of any particular optimization in [10.5, 10.6], the error rates curves presented reveal early changes in slope (BER$\sim 10^{-5}$) that are very pronounced.

A variant of this technique, proposed by Benedetto *et al.* [10.1] made it possible to improve its asymptotic performance. An alternative method to build a TTCM with spectral efficiency $q$ bit/s/Hz involves using two RSC codes with rate $q/(q+1)$ and for each of them to puncture $q/2$ information bits ($q$ is assumed to be even). For each elementary code we thus transmit only half the information bits and all the redundancy bits. The bits at the output of each encoder are associated with a modulation with $2^{(q/2)+1}$ points. The same operation is performed for the two RSC codes, taking care that each systematic bit is transmitted once and only once, so that the resulting turbo code is system-

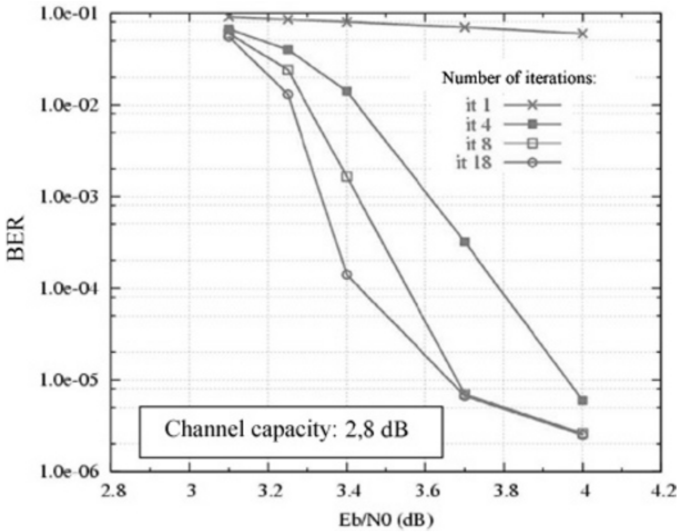



Figure 10.3 – Binary error rate (BER) as a function of the signal to noise ratio $E_b/N_0$ of the 8-PSK TTCM with 8 states using the RSC code of Figure 10.2(a). Transmission over a Gaussian channel. Spectral efficiency $\eta = 2$ bit/s/Hz. Blocks of 10,000 information bits, 5,000 modulated symbols. MAP decoding algorithm. Curves taken from [10.6].

atic. On the other hand, this technique uses interleaving at bit level, and not at symbol level like in the previous approach.

The criterion for optimizing the TTCM proposed in [10.1] is based on maximizing the *effective Euclidean distance*, defined as the minimum Euclidean distance between two encoded sequences whose information sequences have a Hamming weight equal to 2. Figures 10.5 and 10.6 show two examples of TTCMs built on this principle.

The correction performance of these two TTCMs over a Gaussian channel are presented in Figures 10.7 and 10.8. At high and average error rates, they are close to those given by the scheme of Robertson and Wörz; on the other hand, using interleavers operating on the bits rather than on the symbols has made it possible to significantly improve the behaviour at low error rates.

TTCMs lead to excellent correction performance over a Gaussian channel, since they are an *ad hoc* approach to turbo coded modulation. However, they have the main drawback of very limited flexibility: a new code must be defined for each coding rate and each modulation considered. This drawback is cumbersome in any practical system requiring a certain degree of adaptability. On the other hand, although they are a quasi-optimal solution to the problem of coded modulations for the Gaussian channel, their behaviour over fading channels like Rayleigh channels leads to mediocre performance [10.9].

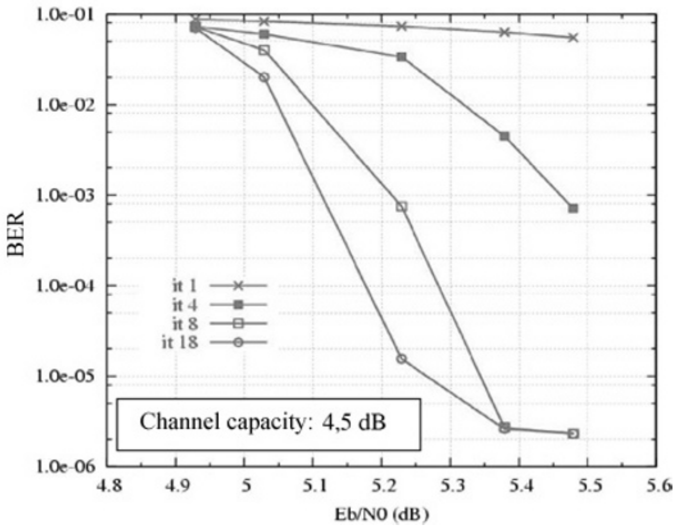


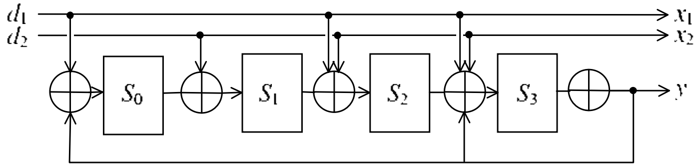

Figure 10.4 – BER as a function of the signal to noise ratio $E_b/N_0$ of the 16-QAM TTCM with 8 states using the RSC code of Figure 10.2(b). Transmission over a Gaussian channel. Spectral efficiency $\eta = 3$ bit/s/Hz. Blocks of 15,000 information bits, 5,000 modulated symbols. MAP decoding algorithm. Curves taken from [10.6].

## 10.2 Pragmatic turbo coded modulation

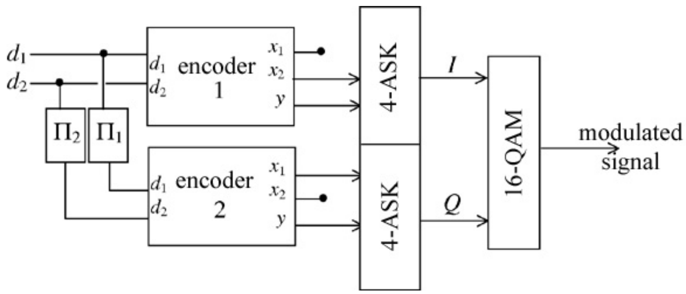
The so-called *pragmatic* approach was chronologically the first implementation. It was introduced by Le Goff *et al.* [10.4] in 1994. This technique takes its name from its similarities with the technique of associating a convolutional code and modulation proposed by Viterbi [10.2] as an alternative solution to Ungerboeck's TCMs. The coding and modulation functions are processed independently, without joint optimization. It uses a "good" turbo code, a bit to signal mapping which minimizes the probability of binary error at the output of the corresponding demapper (Gray coding) and associates the two functions via puncturing and multiplexing to adapt the whole scheme to the spectral efficiency targeted. Figures 10.9 and 10.10 present the general diagram for the principle of the transmitter and the receiver for the pragmatic association of a turbo code and modulation with $Q = 2^q$states.

With this pragmatic approach to turbo coded modulation, the encoder and the decoder used are standard turbo encoders and decoders, identical for all coding rates and modulations considered. If the size of the blocks of data transmitted is variable, simple parametering of the code's permutation function must allow it to adapt to different sizes.

When the targeted coding rate is higher than the natural rate of the turbo code, the puncturing operation enables it to erase, that is to say, not transmit,




(a) Elementary encoder

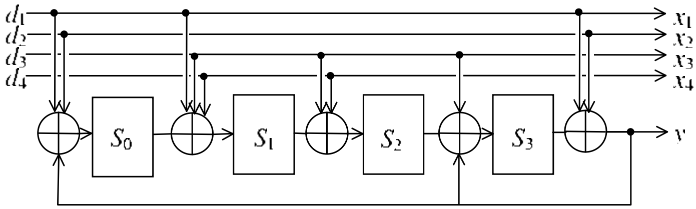



(b) TTCM encoder

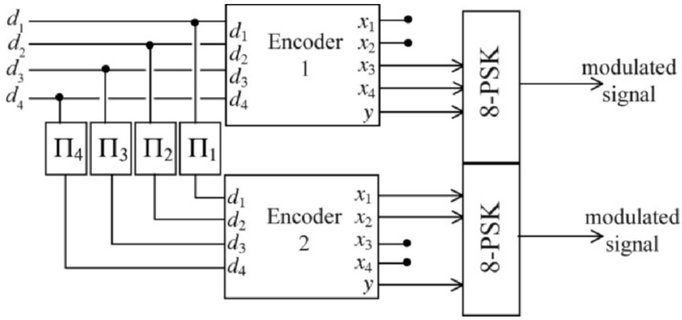
Figure 10.5 – Construction of a 16-QAM TTCM according to the method described in [10.1]. Spectral efficiency $\eta = 2$ bit/s/Hz.

certain coded bits. In practice, for practical reasons of hardware implementation, the puncturing pattern is periodic or quasi-periodic. If possible, only the parity bits are punctured. Indeed, puncturing the systematic bitsleads to a rapid degradation in the decoding convergence threshold, as these bits take part in the process of decoding the two codes, unlike the redundancy bits. When the coding rate is high, a slight puncturing of the data can nevertheless improve the asymptotic behaviour of the system.

The presence of interleaving functions $\Pi^{'}$ and $\Pi^{'-1}$ is justified by the need to decorrelate the data at the input of the turbo decoder. In fact, it is shown in [10.4] that inserting this interleaving has no significant effect on the error rates at the output of the decoder in the case of a transmission on a Gaussian channel. However, in the case of fading channels, the interleaving is necessary as we must prevent bits coming from the same coding instant from belonging to a same symbol transmitted over the channel, so that they will not be affected simultaneously by fading. The studies carried out in this domain [10.9, 10.4, 10.8, 10.2] have shown that in the case of fading channels, the best performance is obtained by using independent interleavers at bit level. This technique is called Bit-Interleaved Coded Modulation (BICM). When the interleaver is placed at modulation symbol level, the order of diversity of the coded modulation is equal to the minimum number of different symbols between two coded modulated sequences. With independent interleavers placed at each output of the encoder, it




(a) Elementary encoder




(b) TTCM encoder

Figure 10.6 – Construction of an 8-PSK TTCM according to the method described in [10.1]. Spectral efficiency $\eta = 2$ bit/s/Hz

.

can ideally reach the Hamming distanceof the code. Consequently, transmission schemes using the BICM principle in practice have better performance on fading channels than TTCMs have.

The code and the modulation not being jointly optimized, unlike a TTCM scheme, we choose binary mapping of the constellation points which minimizes the mean binary error rates at the input of the decoder. When it can be envisaged, Gray encoding satisfies this condition. For simplicity in implementing the modulator and demodulator, in the case of square QAM ($q$ even), the in-phase and in-quadrature axes, $I$ and $Q$, are mapped independently.

In Figure 10.9, the role of the "Multiplexing / symbol composition" block is to distribute the encoded bits, after interleaving for fading channels, into modulation symbols. This block, the meeting point between the code and the modulation, enables a certain level of adjustment of the coded modulation according to the performance targeted. This adjustment is possible since the code and the modulation do not play the same role in relation to all the bits transmitted.

On the one hand, we can distinguish two distinct families of encoded bits at the output of the encoder: systematic bits and redundancy bits. These two families of bits play a different role in the decoding process: the systematic bits, coming directly from the source, are used by the two elementary decoders

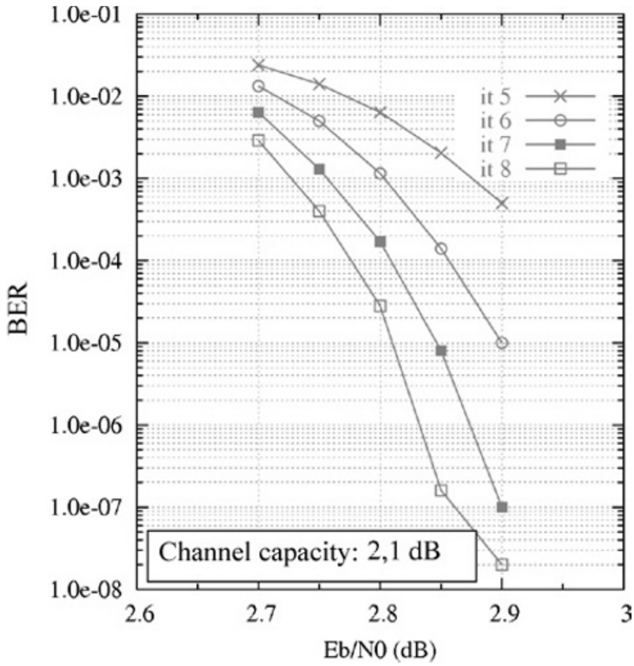



Figure 10.7 – BER as a function of the signal to noise ratio $E_b/N_0$ of the 16-QAM TTCM with 16 states using the RSC code of Figure 10.5. Transmission over a Gaussian channel. Spectral efficiency $\eta = 2$ bit/s/Hz. 2×16,384 information bits. MAP decoding algorithm. Curves taken from [10.1].

at reception whereas the redundancy bits, coming from the two elementary encoders, are used by only one of the two decoders.

On the other hand, the binary elements contained in a modulated symbol are not, in general, all protected identically by the modulation. For example, in the case of PSK or QAM modulation with Gray encoding, only modulations with two or four points offer the same level of protection to all the bits of a same symbol. For higher order modulations, certain bits are better protected than others.

As an illustration, consider a 16-QAM modulation, mapped independently and in an analogue manner on the in-phase and in-quadrature axes by Gray encoding. The projection of this modulation on each of the paths is amplitude shift keying (ASK) with 4 symbols (see Figure 10.11).

We can show that, for a transmission over a Gaussian channel, the error probabilities on the binary positions $s_1$ and $s_2$, given the transmitted symbol ($\pm 3$ or $\pm 1$), are expressed in the form:

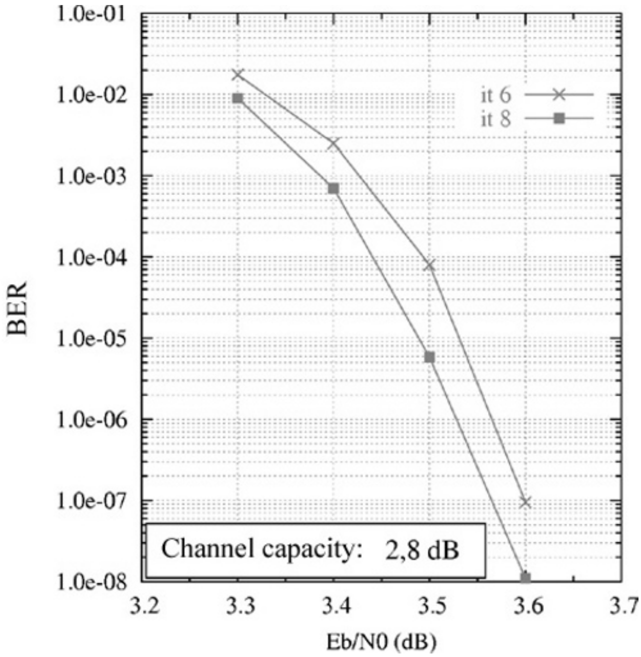



Figure 10.8 – BER as a function of the signal to noise ratio $E_b/N_0$ of the 8-PSK TTCM with 16 states using the RSC code of Figure 10.6. Transmission over a Gaussian channel. Spectral efficiency $\eta = 2$ bit/s/Hz. 4×4,096 information bits. MAP decoding algorithm. Curves taken from [10.3].

$$\begin{aligned}
P_{eb}(s_2 \mid \pm 3) &= \frac{1}{2}\mathrm{erfc}\left(\frac{3}{\sigma\sqrt{2}}\right) \\
P_{eb}(s_2 \mid \pm 1) &= \frac{1}{2}\mathrm{erfc}\left(\frac{1}{\sigma\sqrt{2}}\right) \\
P_{eb}(s_1 \mid \pm 3) &= \frac{1}{2}\mathrm{erfc}\left(\frac{1}{\sigma\sqrt{2}}\right) - \frac{1}{2}\mathrm{erfc}\left(\frac{5}{\sigma\sqrt{2}}\right) \approx \frac{1}{2}\mathrm{erfc}\left(\frac{1}{\sigma\sqrt{2}}\right) \\
P_{eb}(s_1 \mid \pm 1) &= \frac{1}{2}\mathrm{erfc}\left(\frac{3}{\sigma\sqrt{2}}\right) + \frac{1}{2}\mathrm{erfc}\left(\frac{1}{\sigma\sqrt{2}}\right) \approx \frac{1}{2}\mathrm{erfc}\left(\frac{1}{\sigma\sqrt{2}}\right)
\end{aligned} \tag{10.1}$$

where erfc represents the complementary error function and $\sigma^2$ designates the noise variance on the channel. We observe that binary position $s_2$ is on average better protected by the modulation than position $s_1$.

Consequently, it is possible to define several strategies for building modulation symbols by associating as a matter of priority the systematic bits or redun-

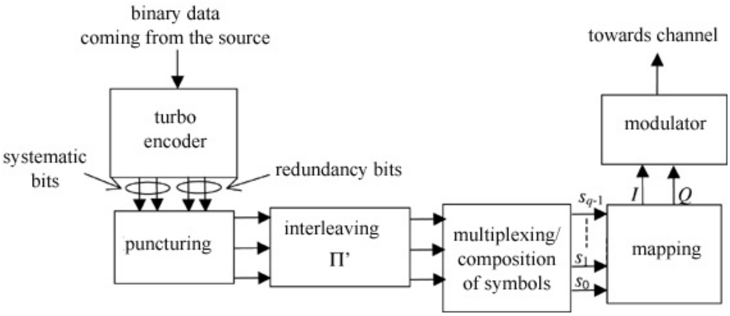



Figure 10.9 – Diagram of the principle of the transmitter in the case of the pragmatic association of a turbo code and modulation with $Q = 2^q$ states.

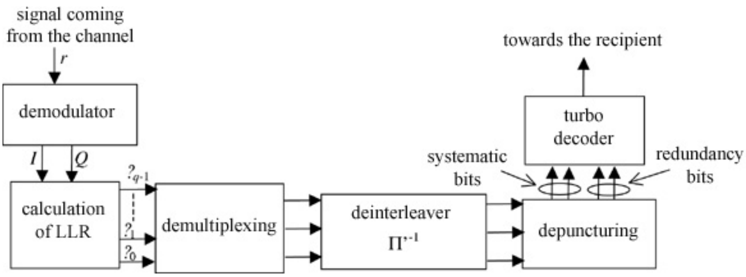



Figure 10.10 – Diagram of the principle of the receiver for the pragmatic turbo coded modulation scheme of Figure 10.9.

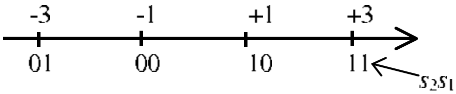



Figure 10.11 – Diagram of the signals of 4-ASK modulation with Gray encoding.

dant bits with the positions that are the best protected by the modulation. Two extreme strategies can thus be defined in all cases:

- so-called "systematic" scheme: the bits best protected by the modulation are associated as a matter of priority with the systematic bits;
- so-called "redundant" scheme: the bits best protected by the modulation are associated as a matter of priority with the redundancy bits.

Modulations of orders higher than 16-QAM offer more than two levels of protection for the different binary positions. 64-QAM modulation, for example, gives three different levels of protection, if the in-phase and in-quadrature axes are mapped independently and in an analogue manner by using a Gray code.

In this case, other schemes, falling in between "systematic" and "redundant" schemes, can be defined.

The reception scheme corresponding to the transmitter of Figure 10.9 is described in Figure 10.10. A standard turbo decoder is used, which requires calculating a weighted estimation of each of the bits contained in the symbols at the output of the demodulator before carrying out the decoding.

The weighted estimation of each bit $s_i$ is obtained by calculating the log likelihood ratio (LLR) defined by:

$$\hat{s}_i = \Lambda(s_i) = \frac{\sigma^2}{2} \ln\left(\frac{\Pr(s_i = 1 \mid r)}{\Pr(s_i = 0 \mid r)}\right) = \frac{\sigma^2}{2} \ln\left(\frac{\Pr(s_i = 1 \mid I, Q)}{\Pr(s_i = 0 \mid I, Q)}\right) \tag{10.2}$$

The sign of the LLR provides the binary decision on $s_i$ (0 if $\Lambda(s_i) \leq 0$, 1 otherwise) and its absolute value represents the weight, that is to say, the reliability, associated with the decision. In the case of a transmission over a Gaussian channel, $\hat{s}_i$ can be written:

$$\hat{s}_i = \frac{\sigma^2}{2} \ln\left(\frac{\sum\limits_{s \in S_{i,1}} \exp\left(-\frac{d_{r,s}^2}{2\sigma^2}\right)}{\sum\limits_{s \in S_{i,0}} \exp\left(-\frac{d_{r,s}^2}{2\sigma^2}\right)}\right) \tag{10.3}$$

where $S_{i,1}$ and $S_{i,0}$ represent the sets of points $s$ of the constellation such that the $i^{th}$ bit $s_i$ is equal to 1 or 0, and $d_{r,s}$ is the Euclidean distance between the received symbol $r$ and the constellation point considered $s$.

In practice, the Max-Log approximation is commonly used to simplify the calculation of the LLRs:

$$\ln(\exp(a) + \exp(b)) \approx \max(a, b) \tag{10.4}$$

and the weighted estimations are calculated as:

$$\hat{s}_i = \frac{1}{4}\left(\min_{s \in S_{i,0}} (d_{r,s}^2) - \min_{s \in S_{i,1}} (d_{r,s}^2)\right). \tag{10.5}$$

We note that it is not necessary to know the noise variance on the channel, when this simplification is used.

The other blocks of the receiver perform the inverse operations of the blocks of Figure 10.9. The depuncturing operation corresponds to the insertion of an LLR equal to 0, this is to say, a zero reliability decision, at the input of the decoder for all the non-transmitted bits.

The pragmatic turbo coded modulation approach enables performance very close to that of TTCMs to be obtained. Figures 10.12 and 10.13 show the performance of two pragmatic turbo coded modulations using the double-binary turbo code with 16 states presented in Section 7.5 for transmission conditions similar to those leading to the curves obtained in Figures 10.3 and 10.4. We observe that after 8 decoding iterations, the performance of the two turbo coded modulation families are equivalent down to BERs from $10^{-4}$ to $10^{-5}$. The better behaviour of the pragmatic solution at lower error rates is due, on the one hand, to the use of 16-state elementary codes and, on the other hand, to the careful design of the turbo code interleaver.

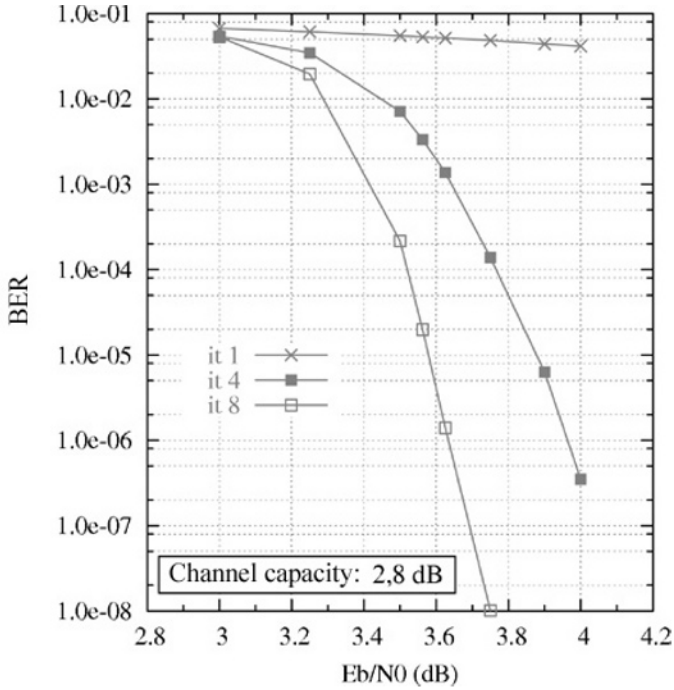



Figure 10.12 – BER as a function of the signal to noise ratio $E_b/N_0$ of pragmatic turbo-coded 8-PSK using a 16-state double-binary code. Transmission over a Gaussian channel. Spectral efficiency $\eta = 2$ bit/s/Hz. Blocks of 10,000 information bits, 5,000 modulated symbols. MAP decoding algorithm. "Systematic" scheme.

The curves of Figure 10.14 show the influence of the strategy of constructing symbols on the performance of turbo coded modulation. They show the behaviour of the association of a 16-state double-binary turbo code and a 16-QAM mapped independently on the in-phase and in-quadrature axes using the Gray code. The two extreme strategies for building the symbols described above were simulated. The size of the blocks, 54 bytes, and the simulated rates 1/2 and 3/4, are representative of concrete applications in the wireless technology sector (IEEE 802.16 standard, *Wireless Metropolitan Area Network*). Figure 10.14 also shows the theoretical limits of the transmission studied. These limits take into

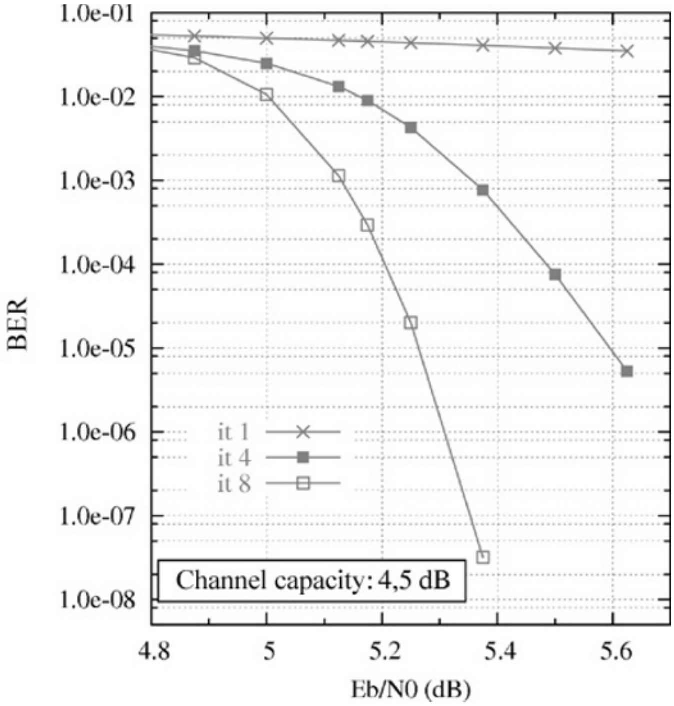



Figure 10.13 – BER as a function of the signal to noise ratio $E_b/N_0$ of pragmatic turbo-coded 16-QAM using a 16-state double-binary code. Transmission over a Gaussian channel. Spectral efficiency $\eta = 3$ bit/s/Hz. Blocks of 15,000 information bits, 5,000 modulated symbols. MAP decoding algorithm. "systematic" scheme.

account the size of the blocks transmitted as well as the packet error rates (PER) targeted. They are obtained from the value of the capacity of the channel, to which we add a correcting term obtained with the help of the so-called "sphere packing" bound, (see Section 3.3).

We observe that at high or average error rates, the convergence of the iterative decoding process is favoured by a better protection of the systematic bits. This result can be explained by the fact that, in the decoding process, each systematic data is used at the input of the two decoders. Consequently, an error on a systematic bit at the output of the channel causes an error at the input of the two elementary decoders, whereas erroneous redundancy only affects the input of one of the two elementary decoders. Consequently, reinforcing the protection of the systematic bits benefits the two elementary decoders simultaneously.

The higher the proportion of redundancy bits transmitted, that is to say, the lower the coding rate, the greater the gap in performance between the two schemes. As an example, at a binary error rate of $10^{-4}$, we observe a gap of 0.9 dB for a coding rate $R = 1/2$, and 0.2 dB for $R = 3/4$.

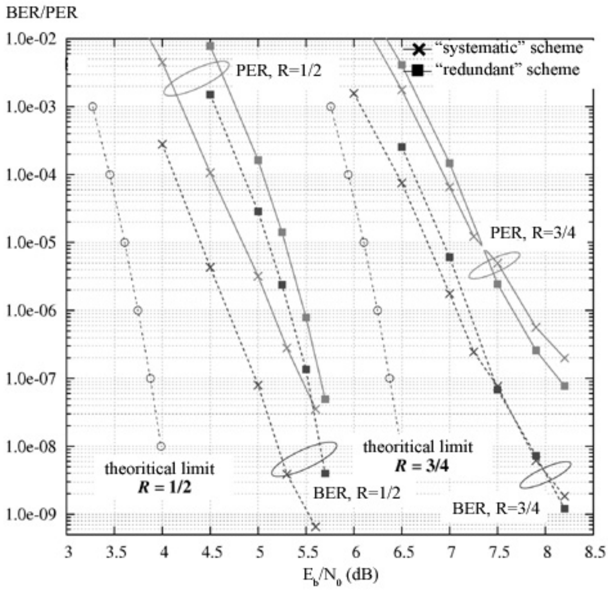



Figure 10.14 – Performance in binary error rate (BER) and packet error rate (PER) of the pragmatic association of a 16-QAM and a 16-state double-binary turbo code, for the transmission of blocks of 54 bytes over a Gaussian channel. Coding rates 1/2 and 3/4. Max-Log-MAP decoding, inputs of the decoder quantized on 6 bits, 8 decoding iterations.

For low and very low error rates, the scheme favouring the protection of the redundancy gives the best performance. This behaviour is difficult to prove by simulation for the lowest rates, as the assumed crossing point of the curves is situated at an error rate that is difficult to obtain by simulation ($PER \approx 10^{-8}$ for $R = 1/2$). The interpretation of this result requires analysis of the erroneous paths in trellises with a high signal to noise ratio. We have observed that, in the majority of cases, the erroneous sequences contain a fairly low number of erroneous systematic bits and a rather high number of erroneous redundancy bits. In other words, the erroneous sequences generally have a low input weight. In particular, the erroneous paths in question mainly correspond to rectangular patterns of errors (see Section 7.3.2). The result, from the point of view of the asymptotic behaviour of turbo coded modulation, is that it is preferable to ensure better protection of the parity bits.

The curves shown in Figure 10.14 were obtained with the help of the simplified Max-Log-MAP decoding algorithm, using data quantized on 6 bits at the input of the decoder. These conditions correspond to a hardware implementation of the decoder. In spite of these constraints, the performance obtained is fairly close to the theoretical limits: only 1 dB at a PER of $10^{-4}$ and 1.5 dB at a PER of $10^{-6}$.

# Bibliography


[10.1] S. Benedetto, D. Divsalar, G. Montorsi, and F. Pollara. Parallel concatenated trellis codes modulation. In *Proceeedings of International Conference on Communications (ICC'96)*, pages 974–978, Dallas, USA, 1996.

[10.2] G. Caire, G. Taricco, and E. Biglieri. Bit-interleaved coded modulation. *IEEE Transactions on Information Theory*, 44(3):927–946, May 1998.

[10.3] C. Fragouli and R. Wesel. Bit vs. symbol interleaving for parallel concatenated trellis coded modulation. In *Proceedings of Global Telecommunications Conference (Globecom'01)*, pages 931–935, San Antonio, USA, Nov. 2001.

[10.4] S. Le Goff, A. Glavieux, and C. Berrou. Turbo codes and high spectral efficiency modulation. In *Proceedings of International Conference on Communications (ICC'94)*, pages 645–649, New Orleans, USA, May 1994.

[10.5] P. Robertson and T. Wörz. Coded modulation scheme employing turbo codes. *Electronics Letters*, 31(2):1546–1547, Aug. 1995.

[10.6] P. Robertson and T. Wörz. Bandwidth-efficient turbo trellis-coded modulation using punctured component codes. *IEEE Journal on Selected Areas in Communications*, 16(2):206–218, Feb. 1998.

[10.7] G. Ungerboeck. Channel coding with mutilevel/phase signals. *IEEE Trans. Info. Theory.*, IT-28(1):55–67, Jan. 1982.

[10.8] A. J. Viterbi, J. K. Wolf, E. Zehavi, and R. Padovani. A pragmatic approach to trellis-coded modulation. *IEEE Communications Magazine*, 27(7):11–19, July 1989.

[10.9] E. Zehavi. 8-psk trellis codes for a rayleigh channel. *IEEE Transactions on Communications*, 40(5):873–884, May 1992.