Turbo Coding and Decoding for Multiple Antenna Channels

Joseph Boutros, Nicolas Gresset, Loïc Brunel
ENST, 46 Rue Barrault, 75013 Paris, FRANCE
Mitsubishi Electric, ITE-TCL, 1 Allée de Beaulieu, 35708 Rennes, FRANCE
Phone: (+33) 1 45 81 76 78, Fax: (+33) 1 45 89 00 20
E-mail: {boutros,gresset,brunel}@enst.fr

***Abstract*— In this paper, we design bit-interleaved coded modulations for multiple antenna channels under iterative detection and decoding. Binary QAM mappings and linear space-time precoders are fine tuned by the mean of a genie method that assumes perfect *a priori* information. We also describe the construction of an interleaver that optimizes the total space-time diversity order for a given error-correcting code. Computer simulations illustrate the universality of bit-interleaved coded modulations with different channel conditions.**

***Index Terms*— space-time codes, multiple antenna channels, iterative decoding.**

I. Introduction

We study bit-interleaved coded modulations (BICM) for digital transmission over multiple-input multiple-output (MIMO) channels. Channel coding techniques for MIMO channels, commonly known as space-time coding, can be classified into four major categories:

- Multi-dimensional trellis coded modulations (TCM) [2][6]. This category includes Ungerboeck-like coded modulations and the simple case of a classical convolutional code where each trellis transition is associated to one channel use.
- Space-time block coding (STBC). The latency of STBC is minimal compared to other techniques. This category includes orthogonal and quasi-orthogonal designs (OD and QOD) [11][15] and the simple technique proposed by Alamouti [9].
- Multilevel coding (MLC) for multiple antennas. Since the original work by Imai and Hirakawa [1][10], it has been demonstrated that MLC can be applied to any type of channels, i.e., scalar and vector channels. In MIMO channels, different levels for coding are defined on QAM symbols fed at the channel input or directly on the binary labels of those symbols.
- Bit-interleaved coded modulations. Combining the original ideas by Zehavi [3][7], Berrou & Glavieux [4], a coded modulation is built by cascading a convolutional code, a pseudo-random interleaver, a QAM symbol mapper and a MIMO channel. The receiver starts by an APP detection of the multiple antenna channel followed by a SISO decoding of the convolutional code. The latter procedure is iterated a finite number of times, where the convolutional code extrinsics are fed back as *a priori* information to the APP detector [12][17].

The paper is organized as follows: Terminology and system model are presented in the next section, this includes the detector structure. Section III describes the genie method and its application to design mappings and space-time precoders. The construction of a new interleaver for universal BICM is given in section IV. Finally, Monte Carlo simulations are found in section V.

II. System model and notations

We consider a MIMO channel with n_t transmit antennas and n_r receive antennas. We limit our study to the case of frequency non-selective channels and coherent detection at the receiver side. The channel is completely defined by its $n_t \times n_r$ complex matrix H. The channel matrix H is assumed to be known to the receiver and unknown to the transmitter. This simple matrix model for a MIMO channel is also valid in the frequency selective case where a multi-carrier modulator (Orthogonal Frequency Division Multiplexing) can remove selectivity. Generally, we distinguish three types of non-selective MIMO channels: 1- Ergodic channel: The random variable H changes at each channel use. 2- Quasi-static channel: The random variable H is constant all along a transmitted codeword. It changes from one codeword to another. 3- Block-fading channel: A channel codeword observes n_c distinct channel states. This case reduces to the first model if n_c is equal to the codeword time length, and to the second model if $n_c = 1$.
A *universal space-time code*, as recently defined by El Gamal, Hammons, Damen [18][20] and other authors, is a space-time code capable of achieving the best diversity order for any value of n_c. We show that a well-designed BICM is a universal space-time code.

A. Structure of the bit-interleaved coded transmitter

Our system model is the following: A binary error-correcting code $\mathcal{C}$ (e.g., a convolutional code), followed by a deterministic interleaver Π, a M-QAM symbol mapper, a rate-1 space-time spreader S (i.e., a linear precoder) and a serial-to-parallel converter. Fig. 1 illustrates the BICM transmitter structure.

Let R_c denote the coding rate of $\mathcal{C}$ and b the information word at the encoder input. The interleaver Π is selected pseudo-randomly with some deterministic constraints as explained in Section IV. The interleaved codeword fed to the QAM mapper is denoted by c. The complex QAM

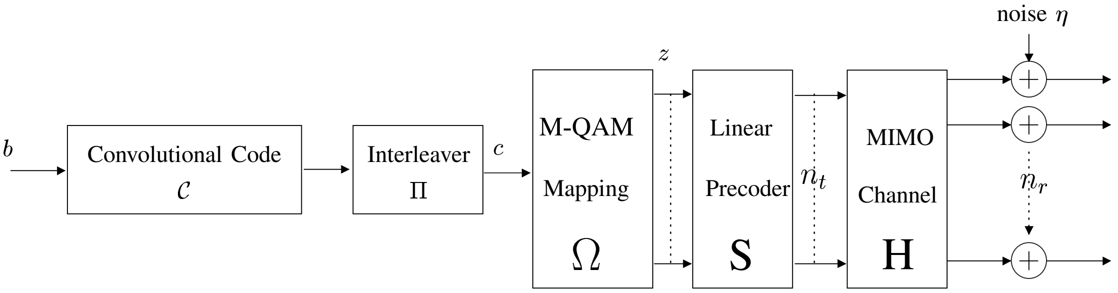


Fig. 1. Bit-interleaver coded mutliple antenna transmitter.

constellation has $M = 2^m$ symbols. At each channel use, the mapper reads $m \times n_t$ coded bits and generates n_t QAM symbols. The linear precoder S spreads the QAM symbols over s time periods. It converts the $n_t \rightarrow n_r$ vector channel into a $N_t \rightarrow N_r$ vector channel, where $N_t = n_t \times s$ and $N_r = n_r \times s$. The $N_t \times N_r$ matrix S reads N_t QAM symbols $z = (z_0, z_1, \dots, z_{N_t-1})$ at the mapper output and generates N_t symbols transmitted during s time periods.
Without space-time spreading ($s = 1$ and S is the identity matrix), the channel path connecting antenna i to antenna j has a complex Gaussian distributed gain h_{ij}, where $H = [h_{ij}]$, $E[h_{ij}] = 0$, $E[|h_{ij}|^2] = 1$, $i = 0 \dots n_t - 1$ and $j = 0 \dots n_r - 1$. Here, the symbol $E[.]$ denotes mathematical expectation. The MIMO channel coefficients h_{ij} are supposed to be statistically independent.

When space-time spreading is applied ($s > 1$), we use the same notation for the extended $N_t \times N_r$ channel matrix

$$H = diag\{ \underbrace{H_1, \dots, H_1}_{s/n_c}, \underbrace{H_2, \dots, H_2}, \dots, \underbrace{H_{n_c}, \dots, H_{n_c}} \} \quad (1)$$

In the above extended block diagonal matrix H, H_t is a $n_t \times n_r$ MIMO matrix corresponding to one channel use at the time period indexed by t. Now, we can write the channel input-output relation:

$$y = x + \eta = zSH + \eta \quad (2)$$

where $y \in \mathbb{C}^{N_r}$ and each receive antenna is perturbed by an additive white complex Gaussian noise η_j, $j = 0 \dots N_r - 1$, with zero mean and variance $2N_0$.

B. Structure of the iterative detector

The spectral efficiency of the transmission system described in the previous sub-section is $R_c \times m \times n_t$ bits per channel use. Let us describe the structure of an iterative a posteriori probability (APP) detector that makes both signal detection and error correction.

The receiver has two main elements as described in Fig. 2: An APP QAM-detector that acts as a soft output equalizer for both the space-time spreader and the MIMO channel, and a SISO decoder for $\mathcal{C}$. An iterative joint detection and decoding process is based on the exchange of soft values between the SISO QAM-detector and the SISO convolutional decoder. The SISO detector computes the extrinsic probabilities $\xi(c_\ell)$ via

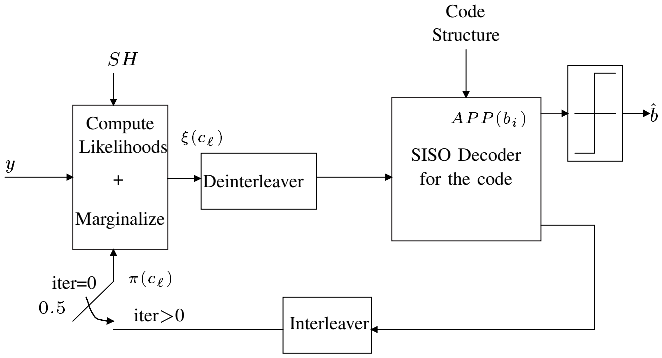


Fig. 2. Iterative APP detection and decoding receiver.

a classical sum product expression including the conditional likelihoods $p(y/z)$ and the *a priori* probabilities $\pi(c_\ell)$ fed back from the SISO decoder. The SISO detector computes the extrinsic information, which corresponds to the extrinsic probability that the j^{th} coded bit equals 1, as given in the following normalized marginalization:

$$\xi(c_\ell) = \frac{\sum_{z' \in \Omega(c_\ell = 1)} \left[\left(e^{-\frac{\|y - z'SH\|^2}{2\sigma^2}} \right) \prod_{r \neq \ell} \pi(c_r) \right]}{\sum_{z \in \Omega} \left[\left(e^{-\frac{\|y - zSH\|^2}{2\sigma^2}} \right) \prod_{r \neq \ell} \pi(c_r) \right]} \quad (3)$$

where Ω is the Cartesian product $(M\text{-QAM})^{N_t}$, i.e., the set of all vectors z generated by the QAM mapper, $|\Omega| = 2^{mN_t}$. The subset $\Omega(c_\ell = 1)$ is restricted to the vectors z where the l^{th} bit is equal to 1. By exploiting the trellis structure of the code, the SISO decoder computes the soft values (a posteriori and extrinsic probabilities) for the coded bits using the Forward-Backward algorithm.

C. Recalling some simple facts in space-time coding

On a general block-fading channel, the performance at the decoder output is function of the so called *diversity order* and *coding gain* [6][18]. Maximizing these two quantities leads to the best space-time code. This is our objective in designing a universal space-time BICM. The coding gain is controlled by the choice of the error-correcting code $\mathcal{C}$ and the binary labeling used in the M-QAM symbol mapper. On the other hand, for a given code $\mathcal{C}$, the diversity order is controlled by the interleaver instance Π and the linear precoder S.

In the literature, authors working on space-time coding studied the design criteria via the pairwise error probability between two transmitted sequences. Mainly, two criteria have been established and are widely used in the space-time coding community:

- **The Hamming distance criterion**: On an ergodic multiple antenna channel, the time diversity order is equal to the minimum columnwise Hamming distance between two distinct codewords. The codewords are viewed as complex matrices with n_t rows.
- **The rank criterion**: On a quasi-static multiple antenna channel, the spatial diversity order is equal to the mimimum rank of the difference matrix between all pairs of distinct codewords.

The pairwise error probability analysis also gives us the expression of the so-called coding gain on a MIMO channel. Except for multi-dimensional rotations on a single antenna Rayleigh fading channel that maximize the product distance [8][19], the coding gain seems to be an uncontrolable parameter in most space-time coding techniques. Fortunately, a BICM will offer us two means for tuning the coding gain: The choice of $\mathcal{C}$ and the mapping design via the genie method as shown in the sequel.

III. The genie method and its applications

We determine below a closed-form expression for the error-rate performance at the detector output in presence of a genie delivering perfect *a priori* information. This expression and the associated figure of merits are called *the genie method*. The genie method enables us to achieve two objectives:

1- Optimize the binary mapping of the QAM constellation, in order to improve the coding gain. The genie method is a very simple tool to compare classical mappings, such as Gray or Ungerboeck mappings, to other optimized mappings.

2- Design a space-time linear precoder that guarantees a maximal diversity order. The linear precoding is made by a modified cyclotomic rotation [8]. The genie method establishes the constraints that must be satisfied by the space-time precoder in order to maximize the transmit diversity measured at the detector output. For simplicity reasons and without loss of generality, it is assumed that $s = 1$ and $S = I$ (identity) in subsections III-A and III-B below.

A. Error performance at the MIMO detector output

Consider the iterative detection and decoding process of the BICM transmitted on the multiple antenna channel as illustrated in Fig. 1 and 2. Assume that extrinsic information associated to such a process is converging towards a limit. The best limit corresponds to the ideal situation where the extrinsic information is perfectly reliable, i.e, $\pi(c_\ell) = c_\ell \in \{0, 1\}$. This is called the *genie* situation. The expression of the detector soft value, when the *a priori* is fed back by a genie, is easily obtained from (3):

$$\xi(c_\ell) = \frac{e^{-\|y-zH\|^2/2N_0}}{e^{-\|y-zH\|^2/2N_0} + e^{-\|y-\bar{z}^\ell H\|^2/2N_0}} \qquad (4)$$

where $\bar{z}^\ell$ is produced by complementing the ℓ^{th} bit in the binary labeling of z. Obviously, from the definition of $\xi(c_\ell)$ in (3), the binary element in the ℓ^{th} position is equal to 1 (respectively 0) in z (respectively $\bar{z}^\ell$). In this case, the system is equivalent to a multidimensional binary modulation BSK with signaling alphabet $\left\{z, \bar{z}^\ell\right\}$ transmitted on a $1 \times n_r$ single input, multiple output (SIMO) channel. We are interested in evaluating the error probability P_e at the detector output when the genie is active. This error probability is directly related to the decision making on $\xi(c_\ell)$. By conditioning on the channel state H and the transmitted QAM vector z, we can write

$$P_{e|H,z} = E_\ell \left[P\left(|\xi(c_\ell) - c_\ell| \geq 0.5\right)\right] \qquad (5)$$

The symbol $E_\ell[.]$ denotes mathematical expectation over the position ℓ of the coded bit. Then, using (4) and (5), we can express P_e with a classical inequality including H, z and η:

$$P_e = E_{H,z,\ell}\left[P\left(\left\|(z - \bar{z}^\ell)H + \eta\right\| \leq \|\eta\|\right)\right] \qquad (6)$$

which is equal to

$$P_e = E_{H,z,\ell}\left[Q\left(\sqrt{\frac{\|(z-\bar{z}^\ell)H\|^2}{4N_0}}\right)\right] \qquad (7)$$

The norm $\left\|(z - \bar{z}^\ell)H\right\|^2$ is calculated from

$$\left\|(z - \bar{z}^\ell)H\right\|^2 = \sum_{u=0}^{n_r-1}\left|\sum_{v=0}^{n_t-1}(z_v - \bar{z_v}^\ell)h_{vu}\right|^2 \qquad (8)$$

We can remark that the performance of the system at the input of the decoder, when the *a priori* feedback at the input of the detector is perfect, is the average probability of the $|\Omega| \times m \times n_t$ equivalent BSKs with distance $d(z, \bar{z}^\ell)$ on a $n_t \times n_r$ MIMO channel. If z and $\bar{z}^\ell$ belong to the same antenna, which is generally the case, this leads to the performance evaluation of a BSK modulation on a $1 \times n_r$ SIMO channel.

The BSK is defined by two equiprobable symbols z and $\bar{z}$, their Euclidean distance is defined by $d(z, \bar{z})$. The received symbol y conditioned on the transmission of z is given by $y = zh + \eta$, where h is a column vector defining the $1 \times n_r$ SIMO channel. The quadratic distance between the two symbols filtered by the channel is $d(z, \bar{z})^2\|h\|^2$. The squared Euclidean norm $\|h\|^2$ has a central χ^2 distribution with degree $2n_r$, the real gaussian random variables have zero mean and variance $1/2$. The noise is white gaussian distributed, the symbols are equiprobable, so we can deduce :

$$P_e = \Phi\left(d(z, \bar{z})^2\right) \qquad (9)$$

The function $\Phi(.)$ is defined as:

$$\Phi\left(d(z, \bar{z})^2\right) = E_h\left[Q\left(\sqrt{\frac{d(z, \bar{z})^2\|h\|^2}{4N_0}}\right)\right] \qquad (10)$$

Similar to a $2n_r$-diversity Rayleigh fading channel (see chap.14 in [13]), the error probability P_e can be calculated in a closed-form expression given in (11).

$$P_e = \Phi\left(d(z,\bar{z})^2\right) = \left(\frac{1-\frac{1}{\sqrt{1+8N_0/d(z,\bar{z})^2}}}{2}\right)^{n_r} \sum_{k=0}^{n_r-1} \binom{n_r-1+k}{k} \left(\frac{1+\frac{1}{\sqrt{1+8N_0/d(z,\bar{z})^2}}}{2}\right)^{k} \tag{11}$$

B. Binary mapping design via the genie method

We consider here the classical situation when a 2-dimensional (complex) labeling is applied on each antenna independently. Let us assume that the coded bit c_ℓ is transmitted on the i-th antenna. Then, $(z-\bar{z}^\ell)$ has only one non-null component in the i^{th} position.
Thanks to (8), we have $\left\|(z-\bar{z}^\ell)H\right\| = d(z,\bar{z}^\ell).\,\|h_i\|$ where h_i is the i-th row of H. In the sequel, the integer i should be considered as a function of the integer ℓ, it can be calculated by $i = \left\lceil \frac{\ell}{m} \right\rceil$. Hence, (11) leads to

$$P_e = \frac{1}{m.n_t.|\Omega|} \sum_{z\in\Omega} \sum_{\ell=1}^{m.n_t} \Phi\left(d(z,\bar{z}^\ell)\right) \tag{12}$$

The asymptotic expression of P_e when $N_0 \to 0$ is:

$$P_e \sim \binom{2n_r-1}{n_r} \left(\frac{2N_0}{\alpha_\Omega}\right)^{n_r} \tag{13}$$

where α_Ω is defined by a harmonic mean:

$$\frac{1}{\alpha_\Omega^{n_r}} = \frac{1}{m.n_t.|\Omega|} \sum_{z\in\Omega} \sum_{\ell=1}^{m.n_t} \frac{1}{d(z,\bar{z}^\ell)^{2n_r}} \tag{14}$$

We can calculate the asymptotic gain of labeling Ω_2 with respect to labeling Ω_1 as follows:

$$Gain_{dB} \sim 10\log_{10}\left(\frac{\alpha_{\Omega_2}}{\alpha_{\Omega_1}}\right) \tag{15}$$

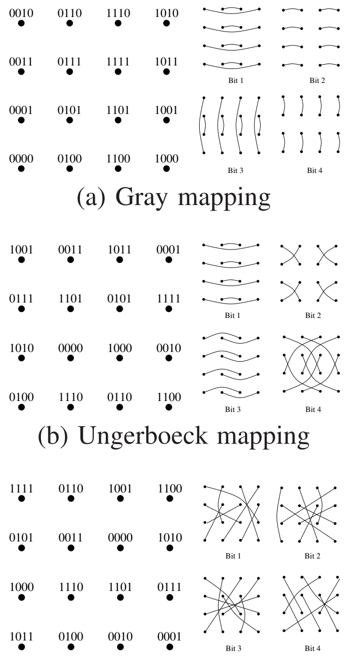
Here, the asymptotic gain only depends on the distance distribution of the equivalent BSKs. We can compare two M-QAM mappings together or a M-QAM mapping with a M-PSK mapping. For example, we found many 16-QAM mappings (selected at random) that exhibit 7dB gain with respect to Gray mapping on a MIMO channel for $n_t = n_r = 2\ldots 8$. One of these binary mappings is shown on Fig. 3. The above results can be considered as a generalization of the single antenna case studied by Chindapol and Ritcey [14].

C. Space-time linear precoding design via the genie method

In the sequel, the $N_t \times N_t$ complex linear spreader S is limited to the case of unitary orthogonal transformations, where $SS^h = S^hS = I$.

1) Space-time codes for the ergodic MIMO channel: The channel changes at each use, i.e., $s = n_c$. The extended channel matrix is $N_t \times N_t$ block diagonal $H = diag\{H_1,\ldots,H_s\}$. If $s=1$ and $S \neq I$, then $H' = SH$ is a different channel instance. Thus, it is trivial to conclude that only the case $s \geq 2$ should be studied.

Let us consider the ℓ-th coded bit c_ℓ of the transmitted vector z when the *a priori* feedback is perfect. As in the previous section, z and $\bar{z}^\ell$ belong to the same QAM symbol indexed by the integer $i = \left\lceil \frac{\ell}{m} \right\rceil$. For a given $z \in \Omega$ and $\ell = 0\ldots m.N_t-1$, let us consider



(a) Gray mapping

(b) Ungerboeck mapping

(c) Optimized mapping

Fig. 3. Three mappings of 16-QAM constellation.

$$\|(z-\bar{z}^\ell)SH\|^2 = d(z,\bar{z}^\ell)^2 \sum_{u=0}^{N_t-1} \left|\sum_{v=0}^{N_t-1} s_{iv}h_{vu}\right|^2 \tag{16}$$

As only modulation symbols with index i differ, and since matrix H is block diagonal, (16) becomes

$$\begin{aligned} &\|(z-\bar{z}^\ell)SH\|^2 = d(z,\bar{z}^\ell)^2 \left\|S_iH\right\|^2 \\ &= d(z,\bar{z}^\ell)^2 \textstyle\sum_{t=0}^{s-1}\sum_{u=0}^{n_r-1} \left|\sum_{v=0}^{n_t-1} s_{i,v+t.n_t} h_{v,u,t}\right|^2 \end{aligned} \tag{17}$$

where $h_{.,.,t}$ denotes the coefficients of the t-th channel matrix H_t and S_i is the i^{th} row of S. Now, $\forall t = 0\ldots s-1$ and $\forall u = 0\ldots n_r-1$,

$$\sum_{v=0}^{n_t-1} s_{i,v+t.n_t} h_{v,u,t} \sim \mathcal{N}\left(0, \sum_{v=0}^{n_t-1} |s_{i,v+t.n_t}|^2\right) \tag{18}$$

$\|S_iH\|^2$ has a generalized chi-square distribution of degree $2N_r$. In the general case, there are s different variance values in the Gaussian variables defining the chi-square law. It is worth noting that two spreading transformations A and B satisfying the equality $\forall t = 0,\ldots,s-1$

$$\sum_{v=0}^{n_t-1} |a_{i,v+t.n_t}|^2 = \sum_{v=0}^{n_t-1} |b_{i,v+t.n_t}|^2 \tag{19}$$

lead to the same genie performance. When the s variances of the Gaussian distributions are not equal, the observed diversity order is not maximal [13]. Based on the latter observation, we derive a first proposition for constructing S:

Proposition 1: If all the norms of the $1/s$-th parts of each row of a $N_t \times N_t$ spreading matrix are equal ($N_t = n_t \cdot s$), the maximum diversity $N_r = n_r \cdot s$ is observed at the output of the detector of a BICM on a $n_t \times n_r$ ergodic MIMO channel.

In this case, when $\forall t = 0, \ldots, s-1, \sum_{v=0}^{n_t-1} |s_{i,v+t.n_t}|^2 = 1/s$, the considered distribution is chi-square with degree $2.N_r$ and parameter $n_t/(2n_s) = 1/(2s)$ (variance of each real dimension). Cyclotomic rotations are an example of transformations satisfying the above norm proposition. Finally, we state the following trivial observation: Under the constraints defined by proposition 1, the total diversity order after precoding is $s.n_r$ which is equivalent to $s.n_r$ receive antennas without precoding.

2) Quasi-static MIMO channel: In this case, $n_c = 1$ and the same fixed $n_t \times n_r$ matrix $H = h_{u,v}$ is applied to the precoding coefficient vector $\{s_{i,t.n_t}, \ldots, s_{i,(t+1).n_t-1}\}$, $\forall\ t = 0, \ldots, s-1$. Then

$$||(z - \bar{z}^\ell)SH||^2 = d(z, \bar{z}^\ell)^2 \sum_{t=0}^{s-1} \sum_{u=0}^{n_r-1} |G_{t,u}|^2 \qquad (20)$$

with $G_{t,u} = \sum_{v=0}^{n_t-1} s_{i,v+t.n_t} h_{v,u}$. The probability law of $||(z - \bar{z}^\ell)SH||^2$ is a correlated chi-square distribution defined by the Gaussian random variables $G_{t,u}$. To get the maximum possible diversity, we have to properly choose the matrix S. First, notice that $\forall (t, t') \in \{0, \ldots, s-1\}^2$ and $\forall (u, u') \in \{0, \ldots, n_r - 1\}^2$

$$u \neq u' \Rightarrow E\left[G_{t,u} G^*_{t',u'}\right] = 0 \qquad (21)$$

because the elements of two different columns of H are independent.

Let us consider the case $u = u'$. We have $E\left[h_{v,u} h^*_{v',u}\right] = 0$, $\forall v \neq v'$. Then

$$\forall (t, t', u), E\left[G_{t,u} G^*_{t',u}\right] = 0 \Leftrightarrow \sum_{v=0}^{n_t-1} s_{i,v+t.n_t} s^*_{i,v+t'.n_t} = 0 \qquad (22)$$

Proposition 2: If the $1/s$-th parts of each row of a $N_t \times N_t$ spreading matrix are orthogonal with equal norm ($N_t = n_t \cdot s$), the maximum diversity $N_r = n_r \cdot s$ is observed at the output of the detector of a BICM on a quasi-static $n_t \times n_r$ MIMO channel.

Of course, the maximal value for the time spreading parameter is $s = n_t$ on the quasi-static fading channel. When the constraints defined by proposition 2 are satisfied, the total diversity order after precoding is $s \cdot n_r$ which is equivalent to $s.n_r$ receive antennas without precoding.

3) Block fading MIMO channel: In order to simplify the notations, $s_{i,w,t,v}$ will designate the coefficient $s_{i,v+(t+sw/n_c)n_t}$ of S. The first index corresponds to the line in the spreading matrix, the second to the part of the line corresponding to the w-th channel state, the third to the t-th time period inside one constant channel block and the fourth to the v-th transmit antenna. The indices i, w, t, v, u belong to the intervals $i \in [0, n_s - 1]$, $w \in [0, c-1]$, $t \in [0, s/c - 1]$, $v \in [0, n_t - 1]$ and $u \in [0, n_t - 1]$. When the linear precoding matrix is affected by n_c different channel states, the same matrix H_w is applied to the vectors $\{s_{i,w,0,0}, \ldots, s_{i,w,s/c-1,n_t-1}\}$. Then

$$||(z - \bar{z}^\ell)SH||^2 = d(z, \bar{z}^\ell)^2 \sum_{w=0}^{n_c-1} \sum_{t=0}^{s/n_c-1} \sum_{u=0}^{n_t-1} |G_{w,t,u}|^2 \qquad (23)$$

where $G_{w,t,u} = \sum_{v=0}^{n_t-1} s_{i,w,t,v} h_{v,u,w}$ is a Gaussian random variable. Again, we note that for two different values of w or u the Gaussian distributions are uncorrelated since they result from the summation of independent Gaussian variables belonging to different channel states or different columns of one channel matrix. The decorrelation criterion becomes, $\forall (t, t', u, w)$:

$$E\left[G_{w,t,u} G^*_{w,t',u}\right] = 0 \Leftrightarrow \sum_{v=0}^{n_t-1} s_{i,w,t,v} s^*_{i,w,v,t'} = 0 \qquad (24)$$

Proposition 3: If the $1/n_c$-th parts of the rows of a $N_t \times N_t$ spreading matrix have the same norm ($N_t = n_t \cdot s$) and if in each $1/n_c$-th part, the n_c/s-th parts are orthogonal together with equal norm, the maximum diversity $N_r = n_r \cdot s$ is observed at the output of the detector of a BICM on a block fading $n_t \times n_r$ MIMO channel with n_c channel states.

The maximal value for the time spreading parameter is $s = n_t \times n_c$. For the construction of a linear precoder on $n_t \times n_r$ block fading channels with n_c different channel states and a time spreading factor equal to s, we can apply a $(N_t/n_c \times N_t/n_c)$ space-time spreading matrix that maximize the diversity for quasi-static sub-channels (proposition 2), and then apply a general $(N_t \times N_t)$ space-time spreading matrix that maximizes the diversity for the ergodic-like channel (proposition 1). As mentioned before, a slight modification of multi-dimensional cyclotomic rotations can be used as a linear spreader in order to improve the diversity order. The algebraic details are omitted for the sake of briefness. The expression of the cyclotmic linear spreader is given in (25), where ϕ is the Euler function.

IV. Interleaver design for universal BICM

The performance after decoding $\mathcal{C}$ depends on the algebraic structure of $\mathcal{C}$ itself and the deterministic instance of the pseudo-random interleaver. Without linear precoding, an upper bound on the maximal diversity order is $n_r \times Min(n_t.n_c, d_{Hmin})$, where d_{Hmin} is the minimum Hamming distance of $\mathcal{C}$. If the BICM interleaver is well chosen and under some conditions (no limitation by the singleton bound), it is possible to obtain the maximal diversity order without space-time precoding. In such conditions, the space-time precoding is seen as another mean for augmenting the diversity order and can be combined to a non-optimized interleaver instance.

$$\begin{array}{l}\forall i \in [0, n_s - 1],\ \forall w \in [0, n_c - 1],\ \forall t \in [0, s/n_c - 1],\ \forall v \in [0, n_t - 1], \\ S_{i,v+(w.s/n_c+t)n_t} = \frac{1}{\sqrt{N_t}} \exp\left(j2\pi\left[i\left(\frac{1}{\phi^{-1}(2N_t)} + \frac{v}{N_t} + \frac{t}{s} + \frac{w}{n_c}\right) + t\left(\frac{1}{\phi^{-1}(2n_t)} + \frac{v}{n_t}\right)\right]\right)\end{array} \quad (25)$$

The interleaver Π should guarantee the minimum columnwise Hamming distance criterion in order to get the temporal diversity n_c or d_{Hmin}, and it should guarantee the rank criterion in order to get the spatial diversity n_t. Such deterministic interleavers exist but are not necessary suited for iterative detection/decoding of a multiple antenna BICM. The factor graph representation [16] of a BICM shows how the information propagates between the detector (channel nodes) and the decoder (subcode nodes) via the sum-product algorithm, see Fig. 4. Iterative APP detection/decoding converges to the exact APP if and only if the graph has no cycles. In our case, for a finite length interleaver, the cycle distribution should be suited for iterative decoding, i.e., a graph distribution with a thin tail as obtained with a pseudo-random interleaver. Unfortunately, the exact influence of the cycles distribution on iterative decoding and the shaping of this distribution is still an open problem. Nevertheless, we designed a special interleaver having both a deterministic and a pseudo-random structure to satisfy the space-time constraint and the factor graph constraint as well. The interleaver presented hereafter is a good interleaver in our sense, but not necessarily the best possible interleaver.

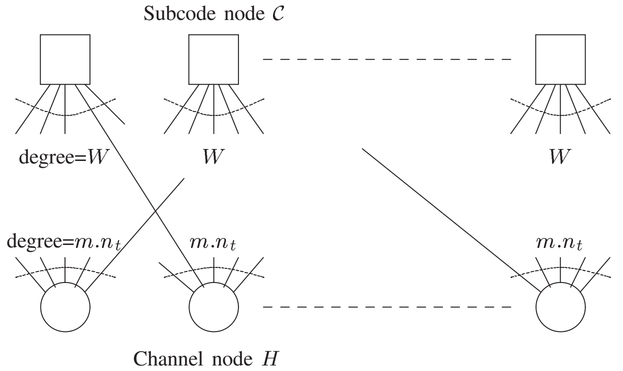


Fig. 4. Factor graph representation of BICM, the window size W is proportionnal to the constraint length of $\mathcal{C}$.

• **Optimized interleaver construction for maximum diversity under iterative decoding**

The different construction phases of the new interleaver are described. The interleaver size is $T.m.n_t$, where T is the time length of a codeword.

Consecutive bits should be mapped on different time periods over all the transmit antennas. To achieve this property, we demultiplex the $T.m.n_t$ coded bits into n_t vectors of length $T.m$. Each one is interleaved separately and mapped diagonally in the space-time domain. This construction will be illustrated for the quasi-static channel ($n_c = 1$) and then extended to the general block fading case. Let us describe the details of the interleaver construction using the example of Fig. 5, with $n_t = 4$ transmit antennas, $n_c = 1$ and $m = 2$ bits per modulation symbol.

The codeword bits are colored in 4 different colors, each one corresponding to a specific transmit antenna. This illustrates the way the codeword is demultiplexed into n_t vectors V_i, $i = 1, \ldots, n_t$, of length $T.m$. Step 1 corresponds to this demultiplexing. Only vector V_1 to be transmitted on antenna 1 is illustrated, containing each fourth bit of the original vector. By this antenna repartition, we already ensure that contiguous bits are transmitted on different antennas.

Each vector V_i of size $T.m$ is then interleaved in step 2 into a vector V_i'. All n_t vectors V_i are interleaved using the same interleaver. In order to ensure a good convergence of the iterative decoding, let us maximize the minimal cycle length in the graph. E.g., an S-random-like interleaver may be used, which guarantees that L_I consecutive bits of V_i before interleaving are placed in different mn_t-length blocks in V_i' after interleaving.

In step 3, we build a $n_t \times T.m$ matrix, each line of which will be transmitted on a different antenna. Let us consider the $m.n_t$ first columns of this matrix. The first row contains the first $m.n_t$ values of the vector V_1' for antenna 1. The second row contains the first $m.n_t$ values of the vector V_2' for antenna 2, shifted by m positions modulo $m.n_t$. Rows 3 and 4 are built from vectors V_3' and V_4' similarly. The T/n_t other $n_t \times m.n_t$ matrices are constructed the same way using the following bits of the n_t vectors V_i'. This results in a cyclic diagonal repartition of the bits in one block of n_t time periods, which guarantees that the bits contained in one symbol period, i.e., in m columns of the final matrix, were originally separated by $(L_I - 1).n_t + 1$ bit positions before interleaving. Moreover it guarantees that $L_I.n_t$ consecutive bits before interleaving are equally distributed on all transmit antennas and mapped on different symbol periods. Practically, the parameter L_I of the S-random-like interleaver should be chosen such that $(L_I - 1).n_t + 1 > L.n$, where L is the constraint length of the code.

Step 4 rewrites the final matrix into a vector of length $T.n_t.m$ assuming that every n_t^{th} bit of the obtained vector will be transmitted on the same antenna (serial to parallel transformation before mapping of m bits into one modulation symbol). The aim of this last step is just to design an interleaver with an input bit stream (vector) and an output bit stream (vector). Of course the structure of the interleaver is related to the way bits will be processed then, i.e., how bits will be distributed onto antennas. From this point of view, the matrix form resulting from step 3 is more explicit.

Let us now focus on the design for block fading channels: $n_c > 1$. The interleaver design is exactly the same as for $n_c = 1$, assuming $n_c.n_t$ transmit antennas. Each vector V_i', $i = 1, \ldots, n_t.n_c$ has length $T.m/n_c$. The first n_t vectors V_i' are transmitted during the T/n_c first time periods, the following n_t vectors V_i' are transmitted during the following T/n_c time periods and so on. Step 4 in Fig. 1 remains the same for the first n_t rows. Then, the following n_t rows are processed the

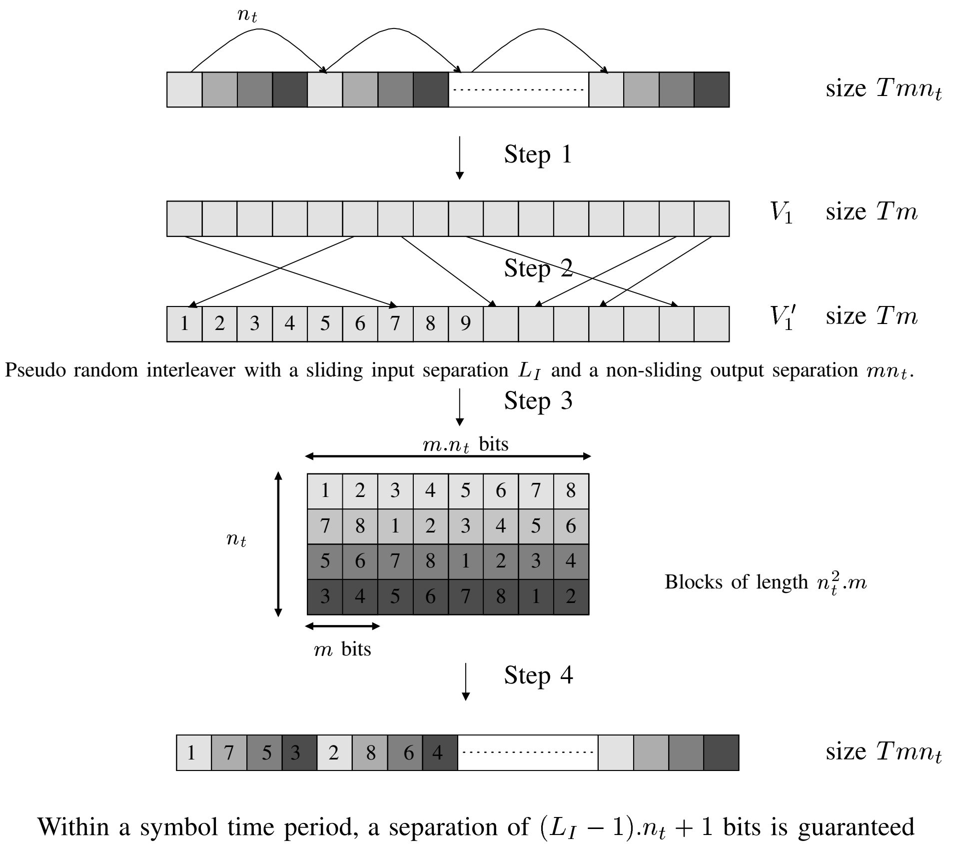


Fig. 5. Illustration of the interleaver construction.

same way, and the obtained vector is written behind the vector for the first n_t rows and so on.

The number of different channel states during a codeword may be known or not. If n_c is unknown, we have to design a universal interleaver that is efficient for any value of n_c. Note that the interleavers designed for a given value of n_c are also well designed for the smaller values of n_c. However, choosing by default a very high value for n_c is not a solution, since it limits the size $T.m/n_c$ of the S-random-like interleaver and thus the performance of iterative decoding. Therefore, we propose to limit n_c to a sufficiently high value by diversity considerations: It is well known [5] that a slight SNR gain is achieved for diversity orders greater than 8. Thus, we can limit the maximal diversity to be collected to $d_{lim} = 16$.

Since the diversity is upper bounded by $n_r.d_{min}$ and limited to d_{lim}, we propose to design an interleaver adapted to a $min(d_{min}, d_{lim}/n_r) \times n_r$ MIMO channel, which should collect a maximum diversity upper bounded by $n_r.min(n_t.n_c, d_{min}, d_{lim}/n_r)$ for any value of n_c. As already said, n_c different channel states is equivalent to a multiplication of the number of antennas by a factor n_c. Thus, the universal interleaver may be constructed as an interleaver designed for a quasi-static channel, assuming that we have $min(d_{min}, d_{lim}/n_r)$ transmit antennas.

V. Computer simulation results

Fig. 6 illustrates the coding gain when comparing an optimized mapping versus a classical Gray mapping for a 16-QAM constellation. The so-called *error-floor region* is below 10^{-5}. In the *waterfall region*, the performance with one simple 4-state rate 1/2 non-recursive non-systematic convolutional code (NRNSC) is very close to that of a rate 1/2 parallel turbo code. The BICM interleaver size is 20000 bits.

Fig. 7 shows the diversity gain with and without linear precoding on a quasi-static channel. The time spreading is $s = 2$. The precoding cyclotomic matrix is 4×4. Three different convolutional codes are displayed with 2 states, 4 states and 16 states respectively.

The performance improvement due to the interleaver structure is presented in Figures 8 and 9. Fig. 8 compares a 4-state and a 16-state convolutional codes on a 2×1 quasi-static channel. For both codes, the pseudo-random interleaver exhibits a poor error rate slope. Fig. 9 displays the frame error rate of a 4-state code on a 4×2 quasi-static channel. Obviously, the interleaver choice is critical. In all cases, the optimized interleaver and the linear precoder lead to a final slope identical to the outage probability one, i.e., maximum achievable diversity.

Finally, the likelihoods involved in the sum-product algorithm on the channel nodes of the BICM factor graph are evaluated exhaustively when the complexity is tractable. Otherwise, if $m \cdot N_t \geq 16$, the evaluation is made by a soft output sphere decoder [21].

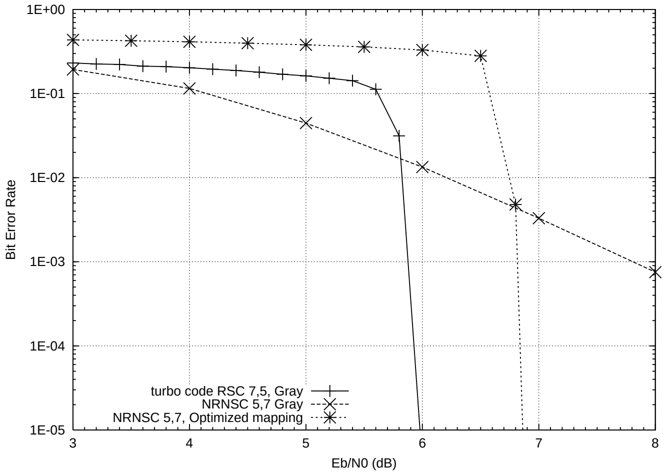


Fig. 6. Performance of a 4-state rate 1/2 convolutional code versus parallel turbo code, 16-QAM constellation, 4x4 MIMO ergodic channel.

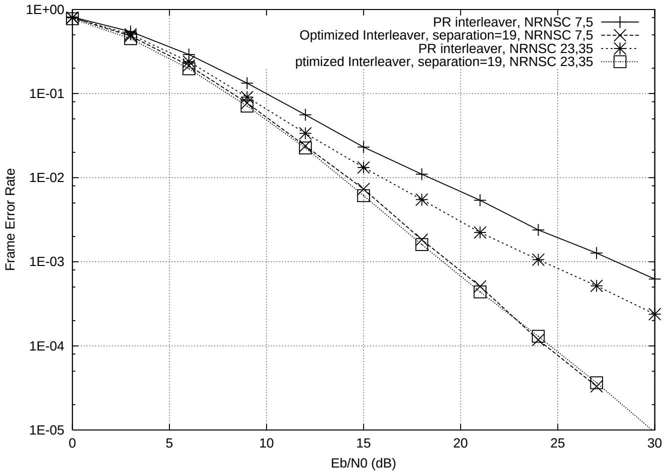


Fig. 8. QPSK on a quasi-static MIMO 2x1, interleaver size=64 bits, 2 decoding iterations.

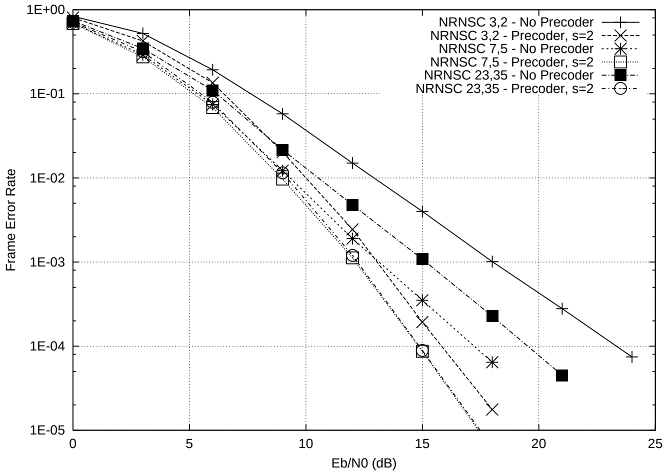


Fig. 7. QPSK on a quasi-static MIMO 2x2, pseudo-random interleaver of size 64 bits, $R_c = 1/2$.

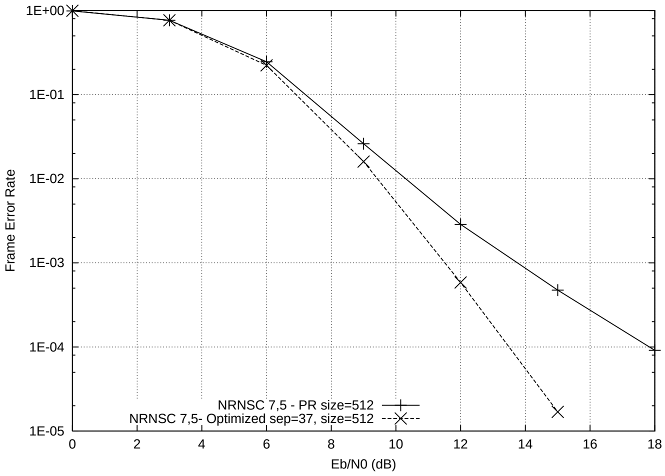


Fig. 9. QPSK on a quasi-static MIMO 4x2, interleaver size=512 bits, 2 decoding iterations.

REFERENCES

[1] H. Imai, S. Hirakawa, “Multilevel coding method using error-correcting codes,” *IEEE Trans. on Information Theory*, vol. 23, pp. 371-377, 1977.

[2] G. Ungerboeck, “Trellis-coded modulation with redundant signal sets, Part II,” *IEEE Communications Magazine*, vol. 25, no. 2, Feb. 1987.

[3] E. Zehavi, “8-PSK trellis codes for a Rayleigh channel,” *IEEE Transactions on Communications*, vol. 40, pp. 873-884, May 1992.

[4] C. Berrou, A. Glavieux, and P. Thitimajshima, ”Near Shannon limit error-correcting coding and decoding: turbo-codes,” *Proceedings of ICC’93*, Geneva, pp. 1064-1070, May 1993.

[5] J. Boutros, E. Viterbo, “New approach for transmission over fading channels,” IEEE Int. Conf. on Univ. Pers. Comm., vol. 1, Sept 1996.

[6] V. Tarokh, N. Seshadri, A.R. Calderbank, “Space-time codes for high data rate wireless communication: Performance criterion and code construction,” *IEEE Trans. on Inf. Theory*, vol. 44, March 1998.

[7] G. Caire, G. Taricco, E. Biglieri, “Bit-interleaved coded modulation,” *IEEE Trans. on Inf. Theory*, vol. 44, no. 3, May 1998.

[8] J. Boutros, E. Viterbo, “Signal space diversity: a power- and bandwidth-efficient diversity technique for the Rayleigh fading channel,” *IEEE Trans. on Inf. Theory*, vol. 44, pp. 1453-1467, July 1998.

[9] S.M. Alamouti, “A simple transmit diversity technique for wireless communications”, *IEEE Journal on Selected Areas in Communications*, vol. 16, pp. 1451 -1458, Oct. 1998.

[10] U. Wachsmann, R.F.H. Fischer, J.B. Huber, “Multilevel codes: theoretical concepts and practical design rules “ *IEEE Trans. Inform. Theory*, vol. 45, pp. 1361-1391, July 1999.

[11] V. Tarokh, H. Jafarkhani, A.R. Calderbank, “Space-time block codes from orthogonal designs,” *IEEE Trans. Inf. Theory*, vol. 45, July 1999.

[12] J. Boutros, F. Boixadera, C. Lamy, “Bit-interleaved coded modulations for multiple-input multiple-output channels,” *IEEE Int. Symp. on Spread Spect. Techniques and Appl.*, vol. 1, pp. 123-126, New Jersey, Sept. 2000.

[13] J. G. Proakis : Digital Communications, 4th edition, *McGraw-Hill*, 2000.

[14] A. Chindapol, J.A. Ritcey,“Design, analysis, and performance evaluation for BICM-ID with square QAM constellations in Rayleigh fading channels,” IEEE Journal on Sel. Areas in Comm., vol. 19, no. 5, pp. 944-957, May 2001.

[15] H. Jafarkhani, “A quasi-orthogonal space-time block code,” *IEEE Trans. on Communications*, vol. 49, PP. 1-4, Jan. 2001.

[16] F. Kschischang, B. Frey, H.-A. Loeliger, “Factor graphs and the sum-product algorithm,” *IEEE Trans. Inf. Theory*, vol. 47, no. 2, Feb. 01.

[17] A. Stefanov, T.M. Duman, “Turbo-coded modulation for systems with transmit and receive antenna diversity over block fading channels: system model, decoding approaches, and practical considerations,” *IEEE Journal on Sel. Areas in Comm.*, vol. 19, pp. 958-968, May 2001.

[18] H. El Gamal, A.R. Hammons, Jr., “On the design of algebraic space-time codes for MIMO block-fading channels,” *IEEE Trans. Inform. Theory*, vol. 49, pp. 151-163, Jan. 2003.

[19] F. Oggier, E. Bayer-Fluckiger, E. Viterbo, “New algebraic constructions of rotated cubic lattice constellations for the rayleigh fading channel,” *IEEE Information Theory Workshop*, pp. 263-266, Paris, March 2003.

[20] H. El Gamal, M.O. Damen, “Universal space-time coding,” *IEEE Trans. Inform. Theory*, vol. 49, pp. 1097-1119, May 2003.

[21] J. Boutros, N. Gresset, L. Brunel, M. Fossorier, “Soft-input soft-output lattice sphere decoder for linear channels,” *IEEE 2003 Global Communications Conference*, San Francisco, Dec. 2003.