Multi-dimensional mappings for iteratively decoded BICM on multiple antenna channels

Nicolas Gresset , *Student Member, IEEE,* Joseph J. Boutros , *Member, IEEE*

and Loïc Brunel , *Member, IEEE*

Abstract

Multi-dimensional binary mappings for bit-interleaved coded modulations on ergodic multiple antenna channels with iterative decoding are presented. After derivation of a closed form expression for the pairwise error probability under ideal maximum likelihood decoding, the design criterion for mapping optimization is established from the maximum likelihood performance of the ideally interleaved channel. It coincides with the figure of merit derived from the genie condition when the iterative receiver converges to perfect a priori information. Multi-dimensional mapping constructions that exhibit high signal to noise ratio gains without increasing the complexity of the a posteriori probability detection are proposed. They allow for a reduced decoding complexity as they achieve near turbo code performance with a single convolutional code.

Index Terms

Binary mapping, bit interleaved coded modulation, multiple antenna channels, pairwise error probability.

N. Gresset and J. Boutros are with the ENST, Paris.

N. Gresset and L. Brunel are with Mitsubishi Electric ITE, Rennes.

I. Introduction

The growing importance of iterative and probabilistic processing of information in communication systems during the last decade has permitted to attain exceptional performance on different kinds of data transmission channels. Graph codes for binary channels have been extensively analyzed [1][2][3][4] and bit-interleaved coded modulations (BICM) for non-binary channels became a widely known standard technique for coded modulations with and without frequency selectivity [5][6][7][8].

Under realistic conditions and without any mild theoretical constraint, the nature of such concatenated systems does not allow us to derive closed-form expressions for the error rate versus the number of decoding iterations. In this paper, we determine the exact pairwise error probability of BICM over multiple antenna channels by assuming an ideal channel interleaver and a maximum likelihood (ML) decoder. Then, we present a tight upper bound on the ideal ML performance of BICM for ergodic multiple antenna channels.

The binary mapping of a signal constellation is an old problem in communication theory. Mappings based on Gray code [9] and Ungerboeck set partitioning [10] are among the most famous binary labelings for coded and uncoded modulations. Multi-dimensional mapping has been extensively studied in the 80's for coded constellations on bandwidth-limited channels, as in [11] for the transmission of fractional bits, in [12][13][14] for trellis-coded multi-dimensional modulations, and in [15] for lattice constellations. More recently, in a parallel work to ours, a multi-dimensional binary mapping and a construction algorithm have been proposed for QPSK on single antenna fading channels [16]. Also, multi-dimensional mappings for multiple antenna BPSK signaling have been described in [17] using a design criterion which is a special case of our figure of merit (see (36) below).

In this paper, a figure of merit for the binary mapping is derived from the ideal ML performance on an ergodic multiple antenna channel. A design criterion based on this figure is applied to the signal constellation to find good mappings suited for space-time coding. Then, it is shown that the mapping figure of merit given by the ML performance is equivalent to the one given by the closed-form expression of the genie performance, related to ideal iterative decoding. The genie method has been previously applied to single antenna fading channels [18] and to multiple antenna channels with mono-dimensional complex mappings [19].

Optimized mappings may be determined either by searching inside a randomly selected list or by applying the binary switching algorithm (BSA) presented in [20][21]. Due to the intractability of the more optimal BSA for large labeling sizes, the first method is used in high complexity systems.

In section II, the system model and notations are presented. In section III, the new tight asymptotic bound for the BICM maximum likelihood performance is computed. It is then compared to genie closed-form performance at the detector output in section IV. In section V, the multi-dimensional mappings are optimized using a figure of merit derived from performance expressions. In section VI, the number of dimensions and the achievable mapping gain are increased by the use of a linear precoder. The convergence behaviour of the optimized mappings is presented in section VII, we show that elementary error-correcting codes (e.g., two-state convolutional code) should be used to achieve the potential gain at relatively low signal to noise ratio (SNR). Finally in section VIII, theoretical results are checked by Monte-Carlo simulations.

II. System Model and Notations

We consider the transmission of a data frame over an ergodic frequency non-selective Rayleigh fading channel with n_t transmit antennas and n_r receive antennas. The channel path connecting antenna i to antenna j during time period k is modeled by a complex Gaussian distributed coefficient h_{kij}, with $E[h_{kij}] = 0$, $E[|h_{kij}|^2] = 1$, $i = 1 \dots n_t$, $j = 1 \dots n_r$ and $k = 1 \dots T$. A frame is transmitted over T time periods. Here, the symbol $E[.]$ denotes mathematical expectation. The multiple input multiple output (MIMO) channel coefficients h_{kij} are supposed to be statistically independent. As usual, the MIMO channel at time k will be represented by its $n_t \times n_r$ matrix $H_k = [h_{kij}]$. Let $\mathcal{H}$ denote the set of channel realizations observed during a frame transmission. As shown in Fig. 1, a binary convolutional encoder with k_c inputs and n_c outputs, followed by a pseudo-random interleaver, generates a codeword c. The length of the convolutional code trellis is L_c branches. The codeword is mapped into n_t-dimensional constellation symbols $z_k \in \Omega \subset \mathbb{C}^{n_t}$, $k = 1 \dots T$, building a frame. In the sequel, the elements of Ω are called *points*. If quadrature amplitude modulation (QAM) is used for digital transmission, then Ω is the Cartesian product $(2^m\text{-QAM})^{n_t}$, i.e., the set of all points z_k generated by the QAM mapper with cardinality $|\Omega| = 2^{mn_t}$. The notation Ω will equally represent the multi-dimensional QAM constellation and its binary labeling of length mn_t bits per point. The channel input-output relation is

$$y_k = z_k H_k + \eta_k \tag{1}$$

where $y_k \in \mathbb{C}^{n_r}$, and η_k is a circularly symmetric complex Gaussian noise with zero mean and covariance matrix $2N_0 I_{n_r}$.

An iterative joint detection and decoding receiver is based on the exchange of soft values between the soft-input soft-output (SISO) multi-dimensional QAM-detector and the SISO convolutional decoder. The SISO detector computes extrinsic probabilities $\xi(c_\ell)$ via a classical sum product expression including the conditional likelihoods $p(y_k|z)$ and the a priori probabilities $\pi(c_\ell)$ fed back from the SISO decoder. The extrinsic probability corresponding to the ℓ^{th} coded bit beeing set to 1 is

$$\xi(c_\ell) = \frac{\sum_{z' \in \Omega(c_\ell=1)} \left[\left(e^{-\frac{\|y_k - z' H_k\|^2}{2N_0}} \right) \prod_{r \neq \ell} \pi(c_r) \right]}{\sum_{z \in \Omega} \left[\left(e^{-\frac{\|y_k - z H_k\|^2}{2N_0}} \right) \prod_{r \neq \ell} \pi(c_r) \right]} \tag{2}$$

The subset $\Omega(c_\ell = 1)$, for $\ell = 1 \dots mn_t$, gathers all vectors z where the ℓ^{th} bit is equal to 1. The SISO decoder computes soft values (a posteriori and extrinsic probabilities) for the coded bits by applying the forward-backward algorithm [23] on the trellis graph of the convolutional code. The information exchange between inputs and outputs of the two received blocks is shown on Fig. 2.

III. Closed form expression for the pairwise error probability under ideal ML decoding

A tight upper bound on the pairwise error probability of error-free decoding for a MIMO-BICM has been given in [24]. It is based on an integral expression that can be evaluated by the Gauss-Chebyshev quadrature [6]. Here, we establish a closed form expression for the exact pairwise error probability on ergodic MIMO channels under maximum likelihood decoding of the BICM and ideal channel interleaving. The mapping design criterion is directly derived from this pairwise error probability expression as shown later in section V. Furthermore, tight union bounds on both frame and bit error rates (FER and BER) will be presented and used to validate the asymptotic signal to noise ratio gain for optimized mappings.

Let $\mathcal{C}$ denote the error correcting code and $\mathcal{C}_E$ the set of Euclidean sequences obtained after applying the QAM mapping Ω and the channel matrix set $\mathcal{H}$ to the binary codewords, i.e., $\mathcal{C}_E$ is the Euclidean code representing the transmitted BICM and including the multiple antenna channel. Consider two codewords $X(c) \in \mathcal{C}_E$ and $X(c') \in \mathcal{C}_E$ with a Hamming distance $w = d_H(c, c')$ between the convolutional codewords c and c'. If we assume ideal channel interleaving, then the w difference positions are spread in space and time over w distinct transmission periods. Clearly, the conditional pairwise error probability $P_{\mathcal{H},w}(X(c) \to X(c'))$ only depends on those w positions. Hence, we will reduce

the notations of $X(c)$ and $X(c')$ to these w time periods. We introduce $X = \{x_1, \ldots, x_w\}$ and $X' = \{x'_1, \ldots, x'_w\}$, where the components x_k and x'_k are points belonging to the set ΩH_k.

Our aim in this section is to compute $P_w(c \to c') = E_{\mathcal{H}}[P_{\mathcal{H},w}(c \to c')]$. The conditional pairwise error probability $P_{\mathcal{H},w}(c \to c')$ is expressed as

$$P_{\mathcal{H},w}(c \to c') = P_{\mathcal{H},w}(X \to X') = P\left(e^{-\sum_{k=1}^{w} \|y_k - x_k\|^2/2N_0} < e^{-\sum_{k=1}^{w} \|y_k - x'_k\|^2/2N_0}\right) \tag{3}$$

For a given set of channel realizations $\mathcal{H}$, a correct pairwise decision is taken by the ML decoder when the log-likelihood ratio $LLR_{\mathcal{H}}$ is positive:

$$\begin{cases} LLR_{\mathcal{H}} & = & \log\left(\frac{e^{-\sum_{k=1}^{w} \|y_k - x_k\|^2/2N_0}}{e^{-\sum_{k=1}^{w} \|y_k - x'_k\|^2/2N_0}}\right) = \frac{\sum_{k=1}^{w} \|y_k - x'_k\|^2 - \|y_k - x_k\|^2}{2N_0} = \sum_{k=1}^{w} LLR_{k,H_k} \\ P_{w,\mathcal{H}}(c \to c') & = & P(LLR_{\mathcal{H}} < 0) = P\left(\sum_{k=1}^{w} LLR_{k,H_k} < 0\right) \end{cases} \tag{4}$$

Thus,

$$P_w(c \to c') = E_{\mathcal{H}}[P(LLR_{\mathcal{H}} < 0)] = E_{\mathcal{H}}\left[\int_{-\infty}^{0} p_{LLR_{\mathcal{H}}}(x)dx\right] = \int_{-\infty}^{0} p_{LLR}(x)dx \tag{5}$$

where $p_{LLR_{\mathcal{H}}}(x)$ is the probability density function of $LLR_{\mathcal{H}}$ and $p_{LLR}(x) = E_{\mathcal{H}}[p_{LLR_{\mathcal{H}}}(x)]$ is the probability density function of $LLR = E_{\mathcal{H}}[LLR_{\mathcal{H}}]$. We will first express the characteristic function $\psi_{LLR}(j\nu)$ of LLR. Since the w random variables LLR_{k,H_k} are independent and the channel is ergodic, using $LLR_{\mathcal{H}} = \sum_{k=1}^{w} LLR_{k,H_k}$, we have

$$\psi_{LLR}(j\nu) = E_{\mathcal{H}}\left[\prod_{k} \psi_{LLR_{k,H_k}}(j\nu)\right] = \prod_{k} \psi_{LLR_k}(j\nu) \tag{6}$$

where $\psi_{LLR_k}(j\nu) = E_{H_k}\left[\psi_{LLR_{k,H_k}}(j\nu)\right]$ and $\psi_{LLR_{k,H_k}}(j\nu)$ is the characteristic function of $p_{LLR_{k,H_k}}(x)$.

Two points are involved in the expression of the partial log-likelihood ratio LLR_{k,H_k}: $x_k = z_k H_k$ and $x'_k = \bar{z}_k^{\ell_k} H_k$, where H_k denotes an instance of the channel matrix set $\mathcal{H}$ at time period k. As ideal interleaving is assumed, the point $z'_k = \bar{z}_k^{\ell_k}$ is obtained by flipping the bit at position ℓ_k in the binary labeling of z_k ($1 \leq \ell_k \leq mn_t$). The squared Euclidean distance between z_k and $\bar{z}^{\ell_k}$ is denoted $d_k^2 = \left\|z_k - \bar{z}_k^{\ell_k}\right\|^2$. The distance spectrum $\{d_k\}$ depends on the modulation type, its size and its binary labeling. For a given 2^m-QAM modulation, non-equivalent labelings lead to non-identical bit error rate performances.

First, we compute the characteristic function of LLR_{k,H_k} for a binary modulation (BSK) defined by two points

$\{z_k,\ \bar{z}_k^{\ell_k}\}$ transmitted over a MIMO channel. The expression of LLR_{k,H_k} is:

$$LLR_{k,H_k} = \frac{1}{2N_0}\left(\left\|y_k - \bar{z}_k^{\ell_k} H_k\right\|^2 - \|y_k - z_k H_k\|^2\right) = \frac{1}{2N_0}\left(d_k^2 R_k + 2\Re\left((z_k - \bar{z}_k^{\ell_k}) H_k \eta_k^*\right)\right) \tag{7}$$

where * is the transpose conjugate and R_k is the squared norm of the vector $\frac{(z_k - \bar{z}_k^{\ell_k})}{d_k} H_k$. If a classical mono-dimensional mapping is used independently on each transmit antenna, the difference vector $z_k - \bar{z}_k^{\ell_k}$ has only one non-null component in the position given by $\lfloor \ell_k/m \rfloor$. However, in order to stay in the general case of multi-dimensional mappings which will be useful in the following, we do not take any assumption on vector $z_k - \bar{z}_k^{\ell_k}$. It can be shown that $\Re\left(\frac{(z_k - \bar{z}_k^{\ell_k})}{N_0} H_k \eta_k^*\right)$ is a Gaussian noise with zero mean and variance $d_k^2 R_k / N_0$. Moreover, since $\frac{(z_k - \bar{z}_k^{\ell_k})}{d_k} H_k$ includes n_r independent identically distributed complex Gaussian random variables with zero mean and unit variance, then R_k has a Chi-square distribution of order $2n_r$. First, notice that the random variable LLR_{k,H_k} is Gaussian distributed.

$$LLR_{k,H_k} \sim \mathcal{N}\left(\frac{d_k^2 R_k}{2N_0}, \frac{d_k^2 R_k}{N_0}\right) \tag{8}$$

The characteristic function of LLR_{k,H_k} is

$$\psi_{LLR_{k,H_k}}(j\nu) = E\left[e^{j\nu LLR_{k,H_k}}\right] = \exp\left(\frac{\nu}{2}\frac{d_k^2 R_k}{N_0}(j-\nu)\right) \tag{9}$$

The mathematical expectation $E_{R_k}[.]$ over R_k is equivalent to the expectation over H_k. Thus,

$$\begin{aligned} \psi_{LLR_k}(j\nu) &= E_{R_k}\left[\psi_{LLR_{k,H_k}}(j\nu)\right] \\ &= \left(1 - \tfrac{d_k^2}{2N_0}\nu(j-\nu)\right)^{-n_r} \\ &= \left(\tfrac{d_k^2}{2N_0}(\nu - ja(d_k))(\nu - jb(d_k))\right)^{-n_r} \end{aligned} \tag{10}$$

where

$$\begin{cases} a(d_k) &= \frac{1}{2}\left(1 + \sqrt{1 + \frac{8N_0}{d_k^2}}\right) \\ b(d_k) &= \frac{1}{2}\left(1 - \sqrt{1 + \frac{8N_0}{d_k^2}}\right) \end{cases} \tag{11}$$

Let D denote the set of all Euclidean distances obtained by flipping one bit in the constellation Ω. Taking n_d as the number of different distances occuring in the sequence $(d_1, d_2, \ldots, d_w)$, we define the set $\Delta = \{\delta_1, \ldots, \delta_{n_d}\} \subset D$ from the sequence $(d_1, d_2, \ldots, d_w) \in \Delta^w \subset D^w$, i.e., the Euclidean distance d_k takes its values from the set Δ. It is clear that $n_d = |\Delta| \leq |D|$. Let the integer λ_k denote the frequency of δ_k in the sequence $(d_1, d_2, \ldots, d_w)$, $\sum_{k=1}^{n_d} \lambda_k = w$ and $\Lambda = \{\lambda_1, \ldots, \lambda_{n_d}\}$.

Using (6), (10) and (11), the averaged characteristic function becomes

$$\begin{aligned}\psi_{LLR}(j\nu) &= \prod_{k=1}^{w}\left(\frac{-d_k^2}{2N_0}(j\nu+a(d_k))(j\nu+b(d_k))\right)^{-n_r} &(12)\\ &= \left(\prod_{k=1}^{w}\left(\frac{-d_k^2}{2N_0}\right)^{-n_r}\right)\left(\prod_{k=1}^{n_d}\left([j\nu+a(\delta_k)][j\nu+b(\delta_k)]\right)^{-n_r\lambda_k}\right) &(13)\\ &= \left(\prod_{k=1}^{w}\left(\frac{-d_k^2}{2N_0}\right)^{-n_r}\right)\left(\prod_{k=-n_d,k\neq 0}^{n_d}[j\nu+\beta_k]^{-n_r\lambda_{|k|}}\right) &(14)\end{aligned}$$

where the poles in the above product are defined by $\beta_{k>0}=a(\delta_k)$, $\beta_{k<0}=b(\delta_{-k})$. To allow derivation of $p_{LLR}(x)$, we now compute the partial fractions expansion of $\psi_{LLR}(j\nu)$:

$$\psi_{LLR}(j\nu)=\prod_{k=1}^{w}\left(\frac{-d_k^2}{2N_0}\right)^{-n_r}\sum_{k=-n_d,k\neq 0}^{n_d}\sum_{i=1}^{n_r\lambda_{|k|}}\frac{\alpha_{k,i}}{(j\nu+\beta_k)^i} \tag{15}$$

The coefficients $\alpha_{\ell,j}$ in (15) can be exactly evaluated from the following identity (series expansion in ϵ):

$$\sum_{i=0}^{n_r\lambda_\ell-1}\alpha_{\ell,n_r\lambda_\ell-i}\epsilon^i+\mathcal{O}(\epsilon^{n_r\lambda_\ell})=\prod_{k=-n_d,k\neq\ell,k\neq 0}^{n_d}\sum_{i=0}^{n_r\lambda_\ell-1}\frac{(-1)^iC^i_{n_r\lambda_{|k|}+i-1}}{(\beta_k-\beta_\ell)^{n_r\lambda_{|k|}+i}}\epsilon^i+\mathcal{O}(\epsilon^{n_r\lambda_\ell}) \tag{16}$$

where $C_n^k=\frac{n!}{k!(n-k)!}$.

From the simple properties $a(\delta_k)-1/2=1/2-b(\delta_k)$ and $\psi_{LLR}(j\nu-1/2)=\psi_{LLR}(-1/2-j\nu)$, we have $\alpha_{-k,i}=(-1)^i\alpha_{k,i}$. Hence, coefficients $\alpha_{k,i}$ are only evaluated for $k>0$. Expression (15) becomes

$$\psi_{LLR}(j\nu)=\prod_{k=1}^{w}\left(\frac{-d_k^2}{2N_0}\right)^{-n_r}\sum_{k=1}^{n_d}\sum_{i=1}^{n_r\lambda_k}\frac{\alpha_{k,i}}{(j\nu+a(\delta_k))^i}+\frac{(-1)^i\alpha_{k,i}}{(j\nu+b(\delta_k))^i} \tag{17}$$

Finally, we get the probability density function of *LLR* by Fourier transform

$$\begin{aligned}p_{LLR}(x) &= \frac{1}{2\pi}\int_{-\infty}^{+\infty}\psi_{LLR}(j\nu)e^{-j\nu x}d\nu &(18)\\ &= \frac{1}{2\pi}\prod_{k=1}^{w}\left(\frac{-d_k^2}{2N_0}\right)^{-n_r}\sum_{k=1}^{n_d}\sum_{i=1}^{n_r\lambda_k}\alpha_{k,i}\left[I_i(x,a(\delta_k))+(-1)^iI_i(x,b(\delta_k))\right] &(19)\end{aligned}$$

and the function $I_i(x,a(\delta_k))$ is defined by

$$I_i(x,a(\delta_k))=\frac{(-x)^{i-1}}{(i-1)!}2\pi\ \mathrm{sgn}(a(\delta_k))e^{a(\delta_k)x}\mathbb{H}(-\mathrm{sgn}(a(\delta_k))x) \tag{20}$$

where $\mathrm{sgn}(x)$ is the sign function, and $\mathbb{H}(x)=\mathbb{I}_{x>0}$ is the Heaviside step function.

The conditional pairwise error probability is $P_w(c \to c') = \int_{-\infty}^{0} p_{LLR}(x)dx$ which yields the closed form expression

$$P_w(c \to c') = P_w(\Delta, \Lambda) = \prod_{k=1}^{w} \left(\frac{-d_k^2}{2N_0} \right)^{-n_r} \sum_{k=1}^{n_d} \sum_{i=1}^{n_r \lambda_k} \frac{\alpha_{k,i}}{a(\delta_k)^i} \tag{21}$$

All sequences $(d_1, \ldots, d_w)$ corresponding to the same (Δ, Λ) yield the same pairwise error probability.

We have expressed the pairwise error probability between two given codewords c and c' such that $d_H(c, c') = w$ and the transmission of $c - c'$ is characterized by (Δ, Λ). We now have to average this probability over all possible pairs (c, c'). First, let us consider the averaged pairwise error probability P_w conditioned on $d_H(c, c') = w$:

$$P_w = E_{c,c'|w} \left[P_w(c \to c') \right] = E_{c,c'|w} \left[P_w(\Delta, \Lambda) \right] \tag{22}$$

Averaging over the pairs (c, c') is equivalent to averaging over (Δ, Λ) thanks to the interleaver. Each pair (Δ, Λ) is representative of $w! / \prod_{i=1}^{n_d} \lambda_i!$ equivalent pairs (Z, Z'), where the w-dimensional Z and Z' vectors are the channel inputs leading to X and X', respectively. As a pair (Z, Z') corresponds to a high number of pairs (c, c'), the complexity of a numerical evaluation is dramatically reduced in practice by performing expectation over the sets Δ and Λ:

$$P_w = E_{\{\Delta, \Lambda\}|w} \left[P_w(\Delta, \Lambda) \right] \tag{23}$$

The frame error rate at the decoder output FER^{dec} is upper bounded by the classical union bound

$$\mathrm{FER}^{dec} \leq E_c \left[\sum_{c' \in \mathcal{C}, c' \neq c} P(c \to c') \right] \tag{24}$$

The input-output transfer function of the convolutional code $\mathcal{C}$ is

$$T(I, Z) = \sum_{w=d_{Hmin}}^{+\infty} \sum_{i}^{+\infty} a_{i,w} I^i Z^w \quad \text{and} \quad T(Z) = T(1, Z) = \sum_{w=d_{Hmin}}^{+\infty} a_w Z^w \tag{25}$$

where $a_{i,w}$ is the number of codewords with an output Hamming weight w and an input Hamming weight i. We can now express the approximation of the maximum likelihood frame error rate and bit error rate of the ideally interleaved BICM transmitted over a multiple antenna channnel:

$$\mathrm{FER}^{dec} \leq \sum_{w=d_{Hmin}}^{+\infty} a_w P_w \tag{26}$$

where P_w is given in (23). Equivalently, we have

$$\mathrm{BER}^{dec} \leq \sum_{j} \sum_{w=d_{Hmin}}^{+\infty} \frac{j}{k_c L_c} a_{j,w} P_w \tag{27}$$

The union bound (UB) for convolutional codes is known to be tight on AWGN channels. Our experimental results will show that the union bound provided by (23) and (26) is also tight on a MIMO ergodic channel.

IV. Genie method and performance

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 toward 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. We will apply (21) to compute the genie performance at the output of the detector. Since only one time period is considered, the temporal subscripts k will be omitted in the following. The expression of the detector soft value, when the a priori is fed back by a genie, is easily obtained from (2):

$$\xi(c_\ell) = \frac{e^{-\|y-zH\|^2/2N_0}}{e^{-\|y-zH\|^2/2N_0} + e^{-\|y-\bar{z}^\ell H\|^2/2N_0}} \tag{28}$$

where $\bar{z}^\ell$ is produced by flipping the ℓ^{th} bit in the binary labeling of z. We are interested in evaluating the error probability BER^{det} 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

$$\mathrm{BER}^{det}_{H,z} = E_\ell \left[P \left(|\xi(c_\ell) - c_\ell| \geq 0.5 \right) \right] \tag{29}$$

The symbol $E_\ell[.]$ denotes the mathematical expectation over the position ℓ of the coded bit. Then, using (28) and (29), we can express BER^{det} as a function of the distance $d = d(z, \bar{z}^\ell)$, averaged over H, z and ℓ:

$$\mathrm{BER}^{det} = E_{H,z,\ell} \left[P_H(z \to \bar{z}^\ell) \right] = E_{z,\ell} \left[\Phi \left(d(z, \bar{z}^\ell) \right) \right] \tag{30}$$

where $\Phi \left(d(z, \bar{z}^\ell \right) = E_H \left[P_H(z \to \bar{z}^\ell) \right]$. We can remark that the performance under genie condition at the detector output, or equivalently at the decoder input, is the average probability of the $|\Omega| m n_t$ equivalent BSKs with distance $d(z, \bar{z}^\ell)$ on a $n_t \times n_r$ MIMO channel. We can directly compute the pairwise error probability from (21) choosing $n_d = w = 1$, $d_1 = \delta_1 = d$, $\lambda_1 = 1$. Finally, we just have to identify the coefficients $\alpha_{1,i}$ from

$$\sum_{i=0}^{n_r-1} \frac{\alpha_{1,n_r-i}}{\varepsilon^{n_r-i}} + \mathcal{O}(1) = \sum_{i=0}^{n_r-1} \frac{(-1)^i C^i_{n_r+i-1}}{(\beta_{-1} - \beta_1)^{n_r+i} \varepsilon^{n_r-i}} + \mathcal{O}(1) \tag{31}$$

using $-\beta_1\beta_{-1} = \frac{2N_0}{d^2}$, we can write the closed form expression of $\Phi(d)$:

$$\Phi(d) = \left(\frac{d^2}{2N_0}\right)^{-n_r} \sum_{k=0}^{n_r-1} \frac{C_{n_r+k-1}^k}{(\beta_1-\beta_{-1})^{n_r+k}\beta_1^{n_r-k}} = \sum_{k=0}^{n_r-1} C_{n_r+k-1}^k \left(\frac{\beta_1}{\beta_1-\beta_{-1}}\right)^k \left(\frac{\beta_{-1}}{\beta_{-1}-\beta_1}\right)^{n_r} \quad (32)$$

$$= \left(\frac{1-\frac{1}{\sqrt{1+8N_0/d^2}}}{2}\right)^{n_r} \sum_{k=0}^{n_r-1} C_{n_r+k-1}^k \left(\frac{1+\frac{1}{\sqrt{1+8N_0/d^2}}}{2}\right)^k \quad (33)$$

which is the result presented in [25, chap. 14].

Finally, the error probability at the detector output is given by

$$\mathrm{BER}^{det} = E_{z,\ell}\left[\Phi\left(d(z,\bar{z}^\ell)\right)\right] = \frac{1}{mn_t|\Omega|}\sum_{z\in\Omega}\sum_{l=1}^{mn_t}\Phi\left(d(z,\bar{z}^\ell)\right) = E_{\{D\}}\left[\Phi(d)\right] \quad (34)$$

V. Multi-dimensional Mapping Optimization

We have presented an approximation of the BICM performance with ideal interleaving and ML decoding. This approximation is a function of the signal to noise ratio, the number of receive antennas and the error-correcting code. Moreover, it mainly depends on the set of distances D given by the binary mapping bit flipping and does not rely on the constellation shape itself. This allows to evaluate the performance of any constellation, even the most unstructured. The performance computation has been processed in the general case of n_t-dimensional distances d_k.

We will first calculate the figure of merit to be optimized for a given n_t-dimensional modulation Ω thanks to its associated distance set D. Then we will apply such an optimization to the classical QAMs and introduce the multi-dimensional mapping concept.

A. Mapping figure of merit

Let us first extract the asymptotic coding gain from the genie performance at the detector output. The asymptotic expression of BER^{det} when $N_0 \to 0$ is

$$\mathrm{BER}^{det} \sim C_{2n_r-1}^{n_r}\frac{(2N_0)^{n_r}}{\mathfrak{F}_\Omega^{det}} \quad (35)$$

where the figure of merit $\mathfrak{F}_\Omega^{det}$ can be computed via [19]

$$\frac{1}{\mathfrak{F}_\Omega^{det}} = E_{\{D\}}\left[d^{-2n_r}\right] = \frac{1}{mn_t|\Omega|}\sum_{z\in\Omega}\sum_{\ell=1}^{mn_t}\frac{1}{d(z,\bar{z}^\ell)^{2n_r}} \quad (36)$$

This figure of merit is equal to the one presented in [18] for $n_t = n_r = 1$. The asymptotic expression of BER^{dec} when $N_0 \to 0$ is

$$\mathrm{BER}_{N_0\to 0}^{dec} = \sum_{j=1}^{+\infty}\frac{j}{k_c L_c}a_{j,d_{Hmin}}C_{2n_r d_{Hmin}-1}^{n_r d_{Hmin}}E_{\{D\}}\left[\prod_{k=1}^{d_{Hmin}}\left(\frac{d_k^2}{2N_0}\right)^{-n_r}\right] \quad (37)$$

Indeed, the error events with Hamming weight greater than d_{Hmin} have higher diversity and negligible contribution to the performance for high signal to noise ratios. The distances in the sequence $(d_1, \ldots, d_{d_{Hmin}})$ are independent random variables thanks to the ideal interleaver. The coding gain is a function of the mapping figure of the merit $\mathfrak{F}_{\Omega}^{dec}$

$$\frac{1}{\mathfrak{F}_{\Omega}^{dec}} = E_{\{D\}}\left[\prod_{k=1}^{d_{Hmin}} d_k^{-2n_r}\right] = \left(E_{\{D\}}\left[d^{-2n_r}\right]\right)^{d_{Hmin}} \tag{38}$$

which leads to $\mathfrak{F}_{\Omega}^{dec} = \left(\mathfrak{F}_{\Omega}^{det}\right)^{d_{Hmin}}$. We notice that optimizing the mapping by maximizing the figure of merit derived from the ML decoding criterion is equivalent to maximizing the figure of merit given by the genie method at the detector output. We can compute the asymptotic gain of labeling Ω_2 with respect to labeling Ω_1 as follows:

$$Gain_{dB} \sim \frac{1}{n_r} 10 \log_{10}\left(\frac{\mathfrak{F}_{\Omega_2}^{det}}{\mathfrak{F}_{\Omega_1}^{det}}\right) \tag{39}$$

The asymptotic gain only depends on the distance distribution of the equivalent BSKs. We can for example compare two QAM mappings together or a QAM mapping with a PSK mapping.

B. Multi-dimensional labelings

When we consider classical mono-dimensional complex labelings, the asymptotic gain optimization is limited by the $m \times n_t$ distances of mono-dimensional complex vectors. Clearly, vectors with more dimensions will lead to higher asymptotic gains. Let us define n_{map} as the number of antennas linked by the labeling.

When performing APP detection, the soft output is computed taking the whole set of transmitted vectors into account. Thus, there is no complexity increase by using a multi-dimensional mapping instead of a mono-dimensional mapping. However, a larger amount of memory is necessary to store the multi-dimensional labelings. When the spectral efficiency is too high, e.g., 4×4 MIMO with 16-QAM input, the exhaustive detector is intractable, and a near-optimum APP detector such as SISO sphere decoder can be used [26]. When using sub-optimum APP detectors such as SISO-MMSE [27], the multiple antenna channel is considered as n_t interfering $1 \times n_r$ SIMO channels, and an exhaustive APP detector is processed on each sub-channel input. In this case, the multi-dimensional mappings cannot be used. The $n_t \times n_r$ MIMO channel can be viewed as n_{part} sub-channels equivalent to $n_t/n_{part} \times n_r$ MIMO channels. We can use a multi-dimensional mapping with $n_{map} \leq n_{part}$, compute an exhaustive detector on each sub-channel and a sub-optimum low complexity detector to separate the n_{part} sub-channels.

C. Mapping optimization

The BICM performance depends on the set of BSK modulations associated to the mapping. For example, the Gray mapping and its associated BSKs are represented in Fig. 3-a. The function $\Phi(d^2)$ defined in (33) is a decreasing function of d^2, this induces that maximizing the BSK distances improves the constellation mapping. Asymptotically, the mapping figure of merit is $\mathfrak{F}_{\Omega}^{det}$ defined in (36). For example, the genie performance of 16-QAM with Gray labeling and minimal Euclidean distance 2.0 is

$$\mathrm{BER}_{Gray}^{det} = \frac{24}{32}\Phi(4) + \frac{8}{32}\Phi(36) \tag{40}$$

The genie performance closed-form expression on MIMO $n_t \times n_r$ channels and the asymptotic gain expression (39) are very useful when designing binary mappings because of the search procedure low complexity. We choose the mapping at random or using an optimization algorithm such as the BSA [20][21]. A mapping is optimized for two parameters: n_r and n_{map}. Indeed, for a given labeling, the asymptotic gain is the same for all n_t. We can numerically determine the asymptotic gain probability distribution of a randomly selected binary mapping, taking the Gray mapping as reference. On Fig. 4, for a 16-QAM constellation, we see the asymptotic gain distribution when $n_r = 1, 2, 4$ and $n_{map} = 1, 2, 4$. We also listed in Table I the mean, variance and maximum value of the asymptotic gain found by our search procedure. We randomly selected a large number of 2^m-QAM mappings, the search is not exhaustive. In the case of $n_r = 1$ and $n_{map} = 1$, the best mapping we found exhibits an asymptotic gain of 7.1 dB. When increasing the mapping number of dimensions ($n_{map} > 1$), it is possible to increase the Euclidean distances of the embedded BSKs. This explains why the statistical mean of asymptotic gain improves for $n_{map} > 1$.

We applied such optimizations to other spectral efficiency values and mapping number of dimensions, the best gains we found with BSA are presented in Table II for 2^m-QAM constellations. In the case of $n_r = 1$ and $n_{map} = 1$, the best mapping found by the BSA exhibits an asymptotic gain of 7.23 dB. This mapping is represented on Fig. 3-b. Unfortunately, the BSA algorithm complexity grows strongly with the global spectral efficiency of the system, that is why we are limited to $mn_{map} < 10$. For $m = 1$ with BSA algorithm, we find the same labelings as constructed in [16][17].

The number of receive antennas n_r has an impact on the figure of merit. Thus, for a same mapping number of dimensions n_{map}, different values of n_r will lead to different optimized mappings. As an example, when n_r tends to infinity, the minimum distance in D will be dominant in the figure of merit expression, as on an AWGN channel, unlike

smaller n_r values.

VI. Increasing the number of dimensions with Space Time precoding

Linear precoding can be used to increase the diversity of systems with a small number of antennas. The symbols of s time periods are grouped together and spread over the transmit antennas and time periods without decreasing the system rate. The linear precoder's matrix S has sn_t rows and columns, where s is called the spreading factor of the linear precoder. A BICM on an ergodic multiple antenna channel exhibits a diversity equal to $d_{Hmin}n_r$. We can increase the observed diversity to $sd_{Hmin}n_r$ using a $sn_t \times sn_t$ complex linear precoder. For example we may use cyclotomic rotations [28][19][29]. If the linear precoder satisfies the norm conditions presented in [19] on an ergodic channel and under a genie condition, maximum precoding gain is obtained and the channel may be seen as a $1 \times sn_r$ SIMO channel. Multi-dimensional mappings designed for sn_r receive antennas may be used without any adaptation. The detection is processed over s time periods. We can use at most a sn_t-dimensional mapping. As shown in section VIII, if $s > 1$, we succeed in enhancing the coding gain via a mapping dimension increase at the cost of detector complexity increase.

VII. Convergence behaviour

We have designed multi-dimensional mappings having large potential gains. Unfortunately, we cannot use such good mappings with a powerful error-correcting code because of convergence problems. Many studies have been made on BICM convergence using exit charts [30] or transfer functions. Most of them conclude that the better the gain at the last iteration, the worst at the first iteration. When considering a joint detection and decoding, the convergence is perfect if the bit error rate at the SISO decoder input in the first iteration is under a given threshold, which corresponds to a SNR value, commonly called waterfall point. The threshold depends on the error-correcting code, and in general, the better the code, the lower the threshold. If the signal to noise ratio is higher than the waterfall point, the system converges to an asymptote after a number of iterations decreasing with the noise level. At very high signal to noise ratio, the mapping gain with respect to Gray mapping is always observed at the output of the error-correcting code. For different mappings, the asymptotes are parallel, their slope is equal to the collected diversity lead by the minimum Hamming distance of the code, the number of receive antennas and the linear precoding factor. If we are interested in a target bit error rate equal to 10^{-5}, we have to find a good compromise between the waterfall and the error floor, as in all iterative processes. In the best case, performance converges to the asymptote exactly at the target error rate. This explains why, when using mappings with high gains, we have to use "bad" error-correcting codes to ensure a good convergence. We illustrate this point on Fig. 5 which represents transfer functions (SNR_{in}^{det},SNR_{out}^{det}) of the detector using a Gaussian

approximation with error probability matching [31] and different mappings. The transfer functions (SNR_{in}^{dec},SNR_{out}^{dec}) of different convolutional codes are also drawn. The transfer functions of recursive systematic convolutional codes show us that the better the code, the higher the slopes. The transfer function of the detector of a 2×2 MIMO channel with SNR = 4.0 dB with QPSK input is also represented with different mappings. The higher the asymptotic gain, the higher the right asymptote, but the lower the left asymptote. We deduce the convergence point searching for the fixed point beginning from the bottom left of the graph. For a given signal to noise ratio, when using multi-dimensional mappings with high asymptotic gain, we should use "bad" error-correcting codes in order to achieve a fixed point close to the right asymptotic value of the detector transfer function. This is equivalent to a perfect convergence to the limit obtained by the genie method.

VIII. Simulation Results

We present some simulation results illustrating the signal to noise ratio gains produced by multi-dimensional labeling under iterative joint detection and decoding. When considering convolutional codes, an exhaustive APP detector computes the soft values delivered to a single SISO decoder: one iteration includes one detection and one forward-backward processing on the convolutional code trellis [23]. When a turbo code is used, one iteration at the receiver side includes one detection, one forward-backward processing on the first convolutional constituent code followed by one forward-backward processing on the second constituent code. A more precise study of the scheduling as in [22] is out of the scope of this paper.

First, Fig. 6 illustrates the error rate of a two-state $(3, 2)_8$ recursive systematic convolutional code (RSC) on a 2×2 MIMO channel with 16-QAM modulation. All situations presented in Fig. 6 correspond to $n_{map} = 1$. Gray mapping is compared to optimized mapping. The latter shows more than 7.4 dB gain with respect to Gray mapping. The three graphs in Fig. 6 show how the simulated error rate quickly converges to the ideal ML bound. The left graph depicts the bit error rate at the decoder output, the middle graph depicts the frame error rate at the decoder output and the right graph depicts the bit error rate at the MIMO detector output.

We now consider a target bit error rate equal to 10^{-5}, usually taken as a reference for wireless data transmissions. The bounds are not drawn anymore.

A convolutional code cascaded with multi-dimensional mappings is compared to a turbo code in Fig. 7. The channel is

2×2 MIMO ergodic with QPSK input. A two-state $(3,2)_8$ non recursive non systematic convolutional code (NRNSC) is combined to mono-dimensional, bi-dimensional and four-dimensional mappings. A parallel turbo code based on an RSC $(7,5)_8$ is cascaded with Gray mapping. Optimized mappings degrade the performance of the turbo code at the first iteration which entails a dramatic signal to noise ratio loss in the waterfall region. Fig. 7 shows that a $(3,2)_8$ convolutional code with four-dimensional mapping ($n_{map} = 4 = sn_t = 2 \times 2$) thanks to the linear precoder [19] performs within 0.5 dB from a rate-1/2 Gray mapped turbo code. The price to pay is the increased detection complexity of the time spreaded ($s = 2$) four-dimensional constellation. The optimized mapping with $n_{map} = 2$ and without linear precoding exhibits excellent error rates above 10^{-3}.

On Fig. 8, we present some simulation results on a 4×4 ergodic MIMO channel with QPSK input and NRNSC $(3,2)_8$. We used mono-, bi- and four-dimensional optimized mappings. We observe that the 0.69 dB (respectively 4.98 dB) expected gain between Gray and mono-dimensional (respectively bi-dimensional) optimized mappings is achieved. When the four-dimensional mapping simulation converges, the asymptote performs lower than 10^{-5}, this is why we measure slightly less than the 7.26 dB expected gain at this BER value. In the latter configuration, the optimal case when the simulation converges to the asymptote exactly at the target BER 10^{-5} is almost achieved. Finally, the system performs as well as the more complex system including turbo code, without increasing the complexity of the detection process. Indeed, in both cases, 20 iterations between the detector and decoder are necessary to achieve the convergence limit and, in each receiver iteration, the turbo decoding is four times more complex than the 2-state convolutional code decoding.

On Fig. 9, we present some simulation results on a 2×2 ergodic MIMO channel with 16-QAM input. When the BER is 10^{-5}, the gain with a mono-dimensional mapping is 7 dB. With a bi-dimensional mapping we achieve 9.1 dB, which is less than the asymptotic 11.12 dB gain because convergence is not reached at 10^{-5}. With high spectral efficiency modulation and a simple NRNSC $(3,2)_8$, we achieve performance within 0.5 dB from the turbo code performance with RSC $(7,5)_8$ constituent codes even on the 2×2 ergodic MIMO channel.

IX. Conclusions

The mapping optimization topic has been extensively discussed for BICM on single antenna channels. In this paper, we have presented an extension of this optimization to multi-dimensional mappings.

We have first presented an exact expression of the pairwise error probability for a BICM over a MIMO channel with

the ideal interleaving assumption. The pairwise error probability is useful to evaluate the BER and FER via a union bound and draw an approximation of the ideal ML performance for moderate and high SNRs. A tangential sphere bound could also be derived to tighten the bound for low SNRs. The union bound has been applied in this paper in the context of mapping optimization. Other straightforward applications of the presented closed form performance could be rotated QAMs or space-time coding. For mapping design, we derived from the union bound a figure of merit and showed that it was equivalent to the one obtained with the more intuitive genie method.

In the case of high spectral efficiency modulations or a large number of transmit antennas, we achieve very high mapping gains and perform close to turbo codes with a single convolutional code, without increasing the optimum or near-optimum APP detector's complexity.

References

[1] "Special issue on codes on graphs and iterative algorithms," *IEEE Trans. on Information Theory*, vol. 47, no. 2, Feb. 2001.

[2] F.R. Kschischang, "Codes defined on graphs," *IEEE Comm. Mag.*, vol. 41, no. 8, pp. 118-125, Aug. 2003.

[3] C. Berrou, "The ten-year-old turbo codes are entering into service," *IEEE Comm. Mag.*, vol. 41, no. 8, pp. 110-116, Aug. 2003.

[4] T. Richardson and R. Urbanke, "The renaissance of Gallager's low-density parity-check codes," *IEEE Comm. Mag.*, vol. 41, no. 8, pp. 126-131, Aug. 2003.

[5] E. Zehavi, "8-PSK trellis codes for a Rayleigh channel," *IEEE Trans. on Communications*, vol. 40, pp. 873-884, May 1992.

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

[7] J. Boutros, F. Boixadera, and C. Lamy, "Bit-interleaved coded modulations for multiple-input multiple-output channels," *Proc. of the IEEE 6th ISSSTA'00*, New Jersey, Sept. 2000.

[8] R. Koetter, A.C. Singer, and M. Tüchler, "Turbo equalization," *IEEE Signal Processing Mag.*, vol. 21, no. 1, pp. 67-80, Jan. 2004.

[9] E. Agrell, J. Lassing, E.G. Ström, and T. Ottosson, "On the optimality of the binary reflected Gray code," *IEEE Trans. on Information Theory*, vol. 50, no. 12, pp. 3170-3182, Dec. 2004.

[10] G. Ungerboeck, "Channel coding with multilevel/phase signals," *IEEE Trans. on Information Theory*, vol. 28, no. 1, pp. 55-67, Jan. 1982.

[11] G. D. Forney, R.G. Gallager, G. R. Lang, F. M. Longstaff, and S. U. Qureshi, "Efficient modulation for band-limited channels," *IEEE Journal on Selected Areas in Communications*, vol. 34, pp. 632-647, Sept. 1984.

[12] G. Ungerboeck, "Trellis-coded modulation with redundant signal sets, part II," *IEEE Comm. Mag.*, vol. 25, no. 2, Feb. 1987.

[13] A. R. Calderbank and N. J. A. Sloane, "New trellis codes based on lattice and cosets," *IEEE Trans. on Information Theory*, vol. 33, no. 2, pp. 177-195, March 1987.

[14] L.F. Wei, "Trellis-coded modulation with multidimensional constellations," *IEEE Trans. on Information Theory*, vol. 33, no. 4, pp. 483-501, July 1987.

[15] G. D. Forney, "Coset codes I: introduction and geometrical classification," *IEEE Trans. on Information Theory*, vol. 34, pp. 1123-1151, Sept. 1988.

[16] N. H. Tran and H. H. Nguyen, "Improving the performance of QPSK BICM-ID by mapping on the hypercube," *Proc. of IEEE VTC Fall'04*, Los Angeles, Sept. 2004.

[17] F. Simoens, H. Wymeersch, and M. Moeneclaey, "Spatial mapping for MIMO systems", *Proc. of IEEE ITW'04*, San Antonio, Oct. 2004.

[18] A. Chindapol and J.A. Ritcey, "Design, analysis, and performance evaluation for BICM-ID with square QAM constellations in Rayleigh fading channels," *IEEE Journal on Selected Areas in Communications*, vol. 19, no. 5, pp. 944-957, May 2001.

[19] J.J. Boutros, N. Gresset, and L. Brunel, "Turbo coding and decoding for multiple antenna channels," *Proc. of the 3rd International Symposium on Turbo Codes'03*, Brest, Sept. 2003. Available at http://www.comelec.enst.fr/∼boutros/coding.

[20] K. Zeger and A. Gersho, "Pseudo-Gray Coding," *IEEE Trans. on Communications*, vol. 38, pp. 2147-2158, Dec. 1990.

[21] F. Schreckenbach, N. Gortz, J. Hagenauer, and G. Bauch, "Optimized symbol mappings for bit-interleaved coded modulation with iterative decoding," *Proc. of the IEEE GLOBECOM'03*, San Francisco, Dec. 2003.

[22] F. Brännström, L. K. Rasmussen and A. Grant, "Optimal scheduling for iterative decoding", *Proc. of IEEE ISIT'03*, Yokohama, Japan, July 2003.

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

[24] Y. Huang and J.A. Ritcey, "Tight BER bounds for iteratively decoded bit-interleaved space-time coded modulation," *IEEE Comm. Letters*, vol. 8, no. 3, pp. 153-155, March 2004.

[25] J.G. Proakis, "Digital Communications," 3rd edition, McGraw-Hill, New York, 1995.

[26] J. Boutros , N Gresset, L. Brunel, and M. Fossorier, "Soft-input soft-output lattice sphere decoder for linear channels," *Proc. of the IEEE GLOBECOM'03*, San Francisco, Dec. 2003.

[27] A. Dejonghe and L. Vanderdorpe, "Turbo-equalization for multilevel modulation: an efficient low-complexity scheme," *Proc. of the IEEE ICC'02*, New York, vol. 25, no. 1, pp. 1863-1867, April 2002.

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

[29] N. Gresset, J. Boutros, and L. Brunel, "Optimal linear precoding for BICM over MIMO channels," *Proc. of the IEEE ISIT'04*, Chicago, June 2004.

[30] S. Ten Brink, "Convergence of iterative decoding", *Electronics Letters*, vol. 35, no. 10, pp. 806-808 ,May 1999.

[31] J.J. Boutros, "Iterative APP decoding of linear transforms", *Proc. of the ISITA'00*, Hawai, November 2000.

Nicolas Gesset was born in Enghien-les-bains, France, in 1978. He received the Ph.D. degree from Ecole Nationale Supérieure des Télécommuncations (ENST), Paris, France, in 2004. He is now with France Telecom R&D, Issy les Moulineaux, France. His research interests are channel coding, multiple-antenna techniques and code division mutliple access.

Joseph J. Boutros received the electrical engineering degree in 1992 and the Ph.D. degree in 1996 from the Ecole Nationale Superieure des Telecommunications (ENST), Paris, France. Since September 1996, he is with the Communications and Electronics Department at ENST as an Associate Professor. His fields of interest are codes on graphs, iterative decoding, joint source-channel coding, space-time coding and lattice sphere packings.

Loïc Brunel was born in Antony, France, in 1973. He received the electrical engineering and Ph.D. degrees from Ecole Nationale Supérieure des Télécommuncations (ENST), Paris, France, in 1996 and 1999, respectively. Since 2000, he has been a Research Engineer in the Telecommunication Laboratory, Mitsubishi Electric ITE, Rennes, France. His research interests are multi-carrier transmissions, code division multiple access, channel coding and multiple-antenna techniques.

	Mean	Variance	Max gain (dB) at random	Max gain (dB) with BSA alg
$n_r = 1$, $n_{map} = 1$	3.15	0.35	7.10	7.23
$n_r = 2$, $n_{map} = 1$	2.40	0.29	7.27	7.42
$n_r = 4$, $n_{map} = 1$	1.43	0.13	7.15	7.36
$n_r = 1$, $n_{map} = 2$	6.75	0.02	7.48	10.68
$n_r = 2$, $n_{map} = 2$	5.65	0.04	6.80	11.12
$n_r = 4$, $n_{map} = 2$	3.59	0.04	5.01	10.98
$n_r = 1$, $n_{map} = 4$	10.97	0.01	10.99	/
$n_r = 2$, $n_{map} = 4$	10.67	0.01	10.71	/
$n_r = 4$, $n_{map} = 4$	8.33	0.02	8.57	/

TABLE I
STATISTICS OF 16-QAM OPTIMIZED MAPPINGS

	$m = 1$	$m = 2$	$m = 4$	$m = 6$
$n_r = 1$, $n_{map} = 1$	0.00	1.25	7.23	12.62
$n_r = 1$, $n_{map} = 2$	1.25	5.05	10.68	/
$n_r = 1$, $n_{map} = 3$	3.55	6.52	/	/
$n_r = 1$, $n_{map} = 4$	5.05	/	/	/
$n_r = 2$, $n_{map} = 1$	0.00	1.02	7.42	12.97
$n_r = 2$, $n_{map} = 2$	1.02	5.02	11.12	/
$n_r = 2$, $n_{map} = 3$	3.46	6.24	/	/
$n_r = 2$, $n_{map} = 4$	5.02	/	/	/
$n_r = 4$, $n_{map} = 1$	0.00	0.69	7.36	12.81
$n_r = 4$, $n_{map} = 2$	0.69	4.98	10.98	/
$n_r = 4$, $n_{map} = 3$	3.35	6.16	/	/
$n_r = 4$, $n_{map} = 4$	4.98	7.26	/	/

TABLE II
BEST FOUND ASYMPTOTIC GAINS (IN DB) WITH RESPECT TO GRAY MAPPING FOR 2^m-QAM CONSTELLATIONS AND n_{map} DIMENSIONS

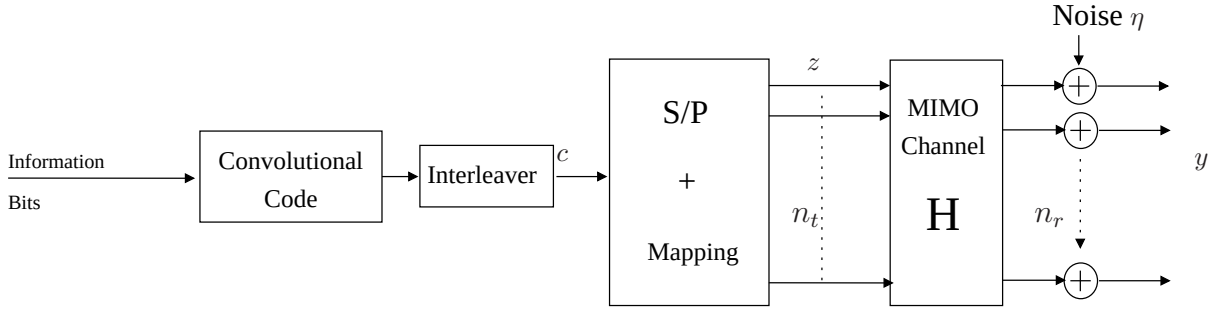

Fig. 1. BICM on the MIMO channel.

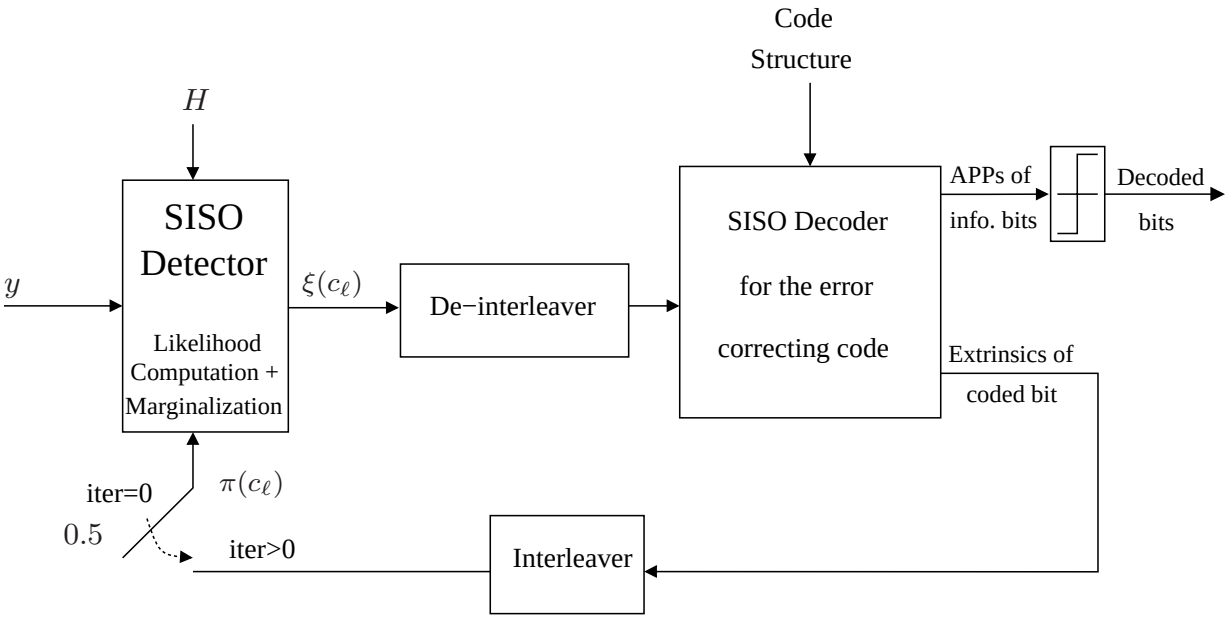

Fig. 2. Joint detection and decoding of a BICM.

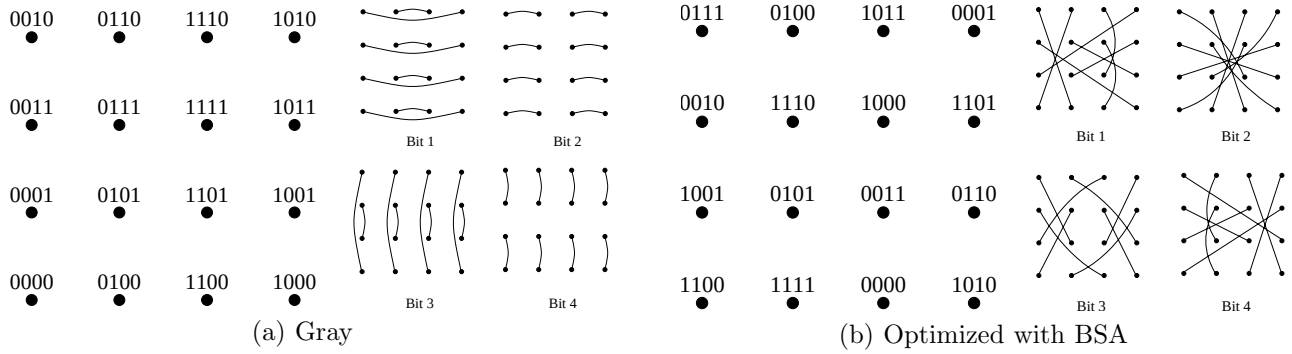

Fig. 3. Mappings of 16-QAM constellation.

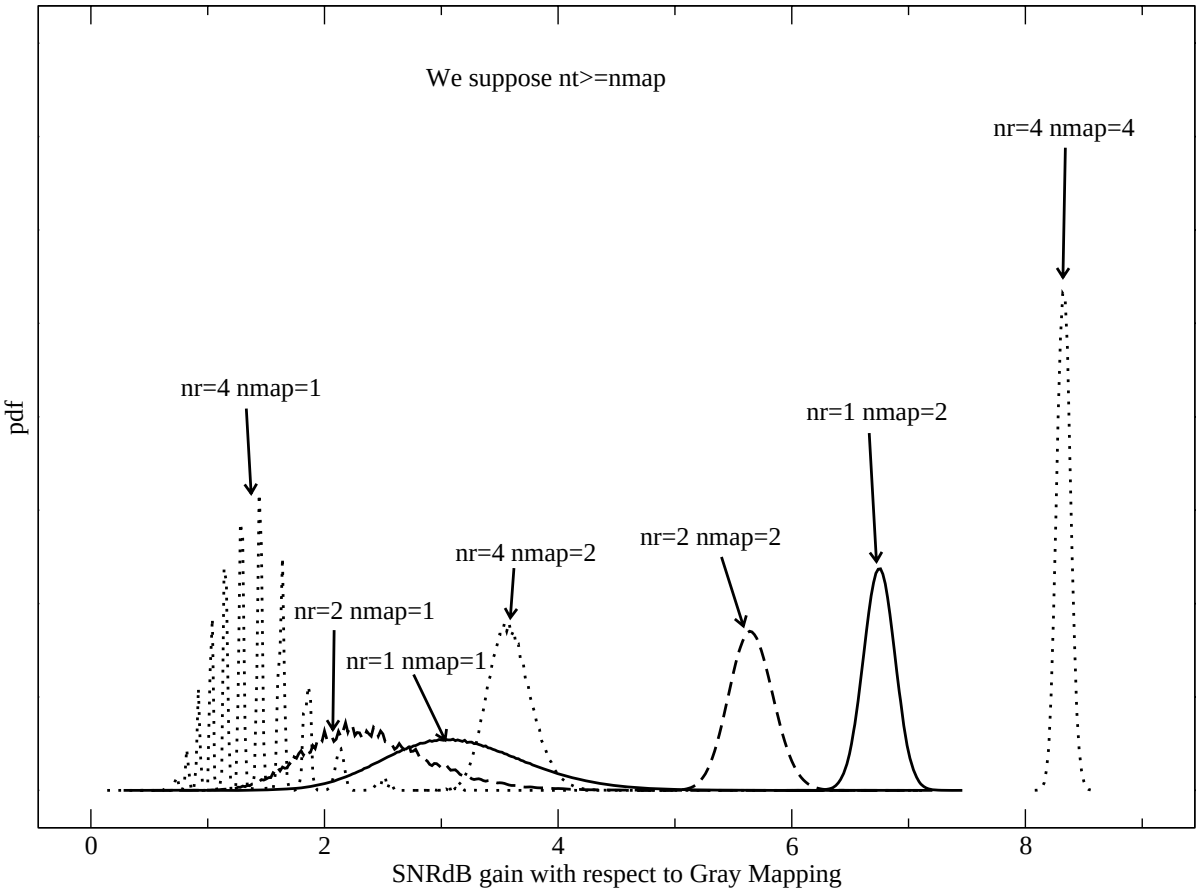

Fig. 4. Asymptotic gain distribution of random 16-QAM mapping with respect to Gray mapping.

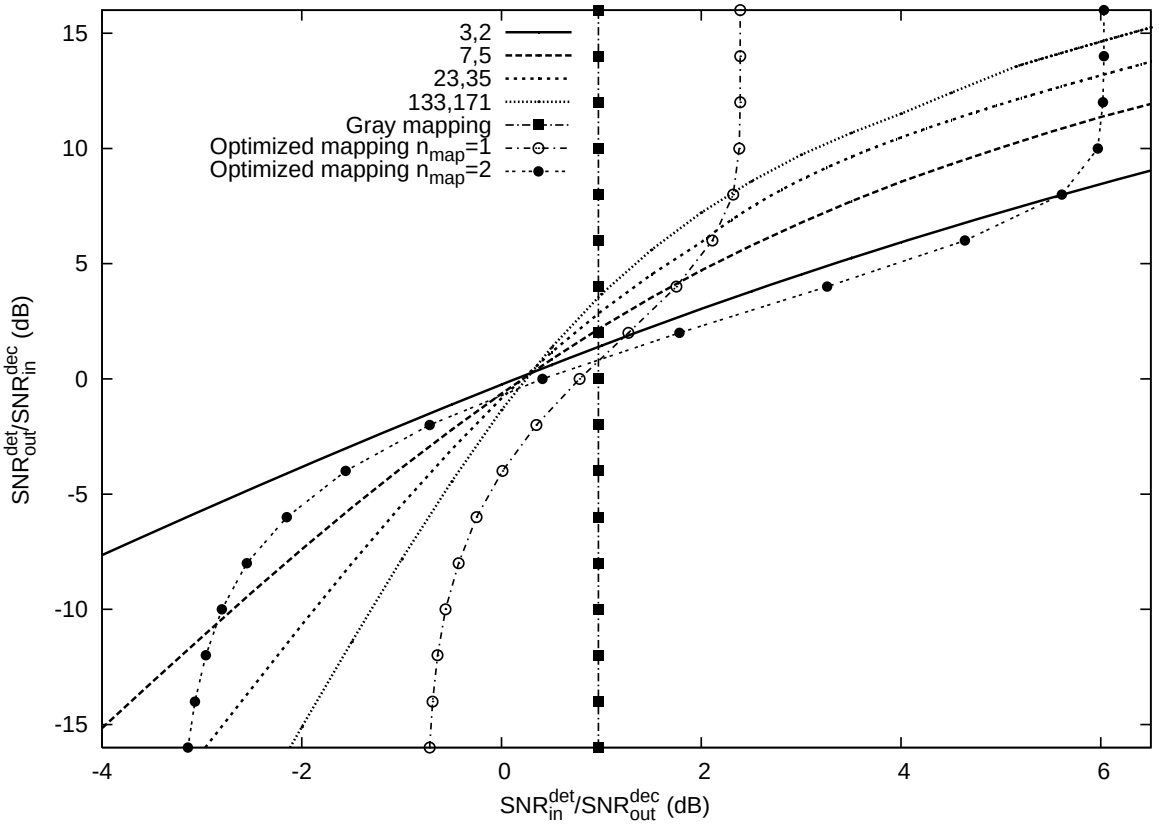

Fig. 5. Transfer function of RSC codes and QPSK multi-dimensional mappings, $n_t = 2$, $n_r = 2$, $SNR = 4.0$ dB

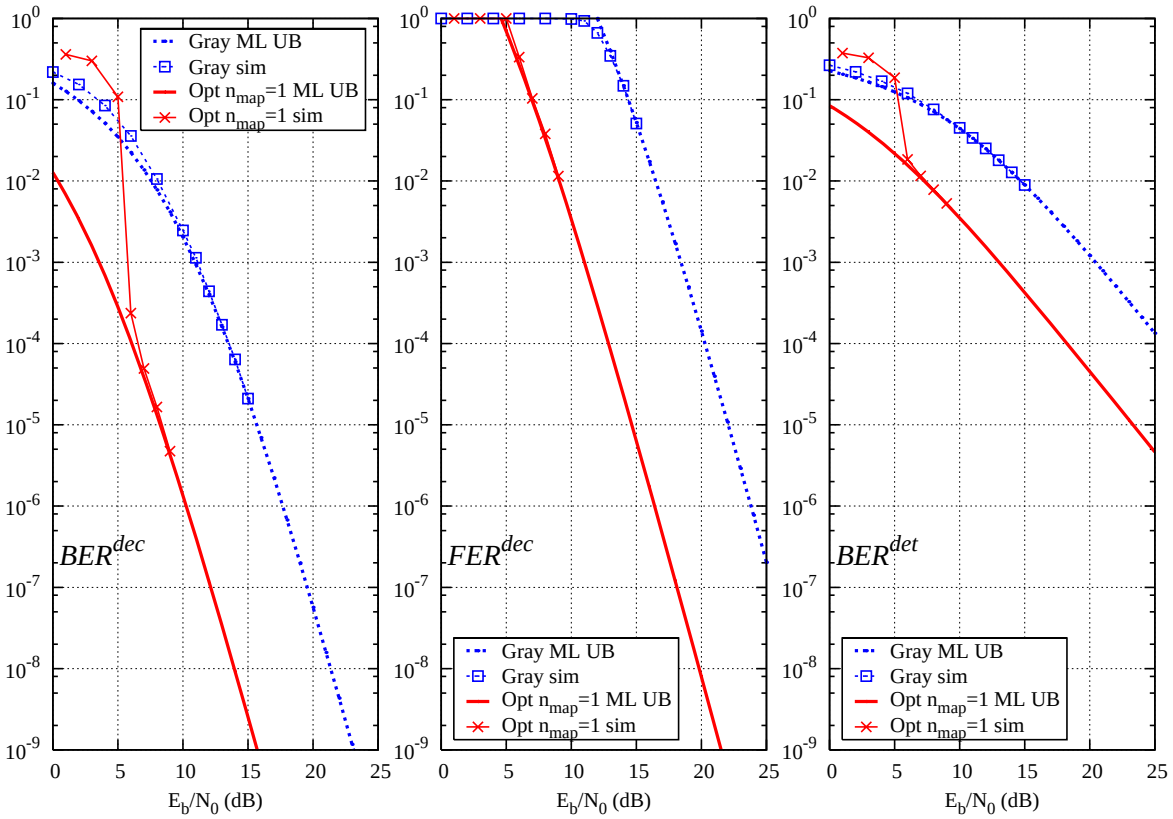

Fig. 6. Ergodic 2×2 MIMO channel, interleaver size is 10000 bits, 2-state $(3,2)_8$ convolutional code, 16-QAM modulation, 10 decoding iterations. ML upper bound is denoted by "ML UB" and Monte Carlo simulation is denoted by "sim" in the captions.

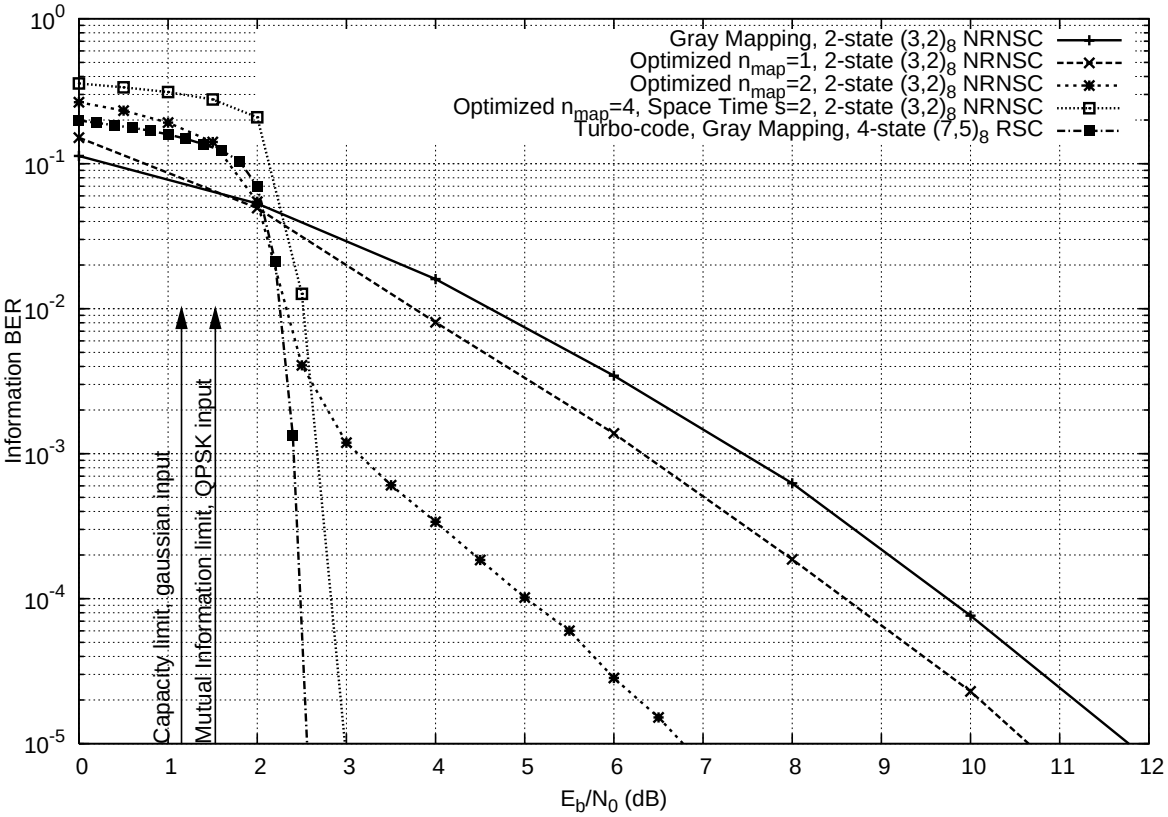

Fig. 7. Ergodic 2×2 MIMO channel, interleaver size is 9000 bits, rate-1/2 NRNSC and turbo code, QPSK modulation, 20 decoding iterations.

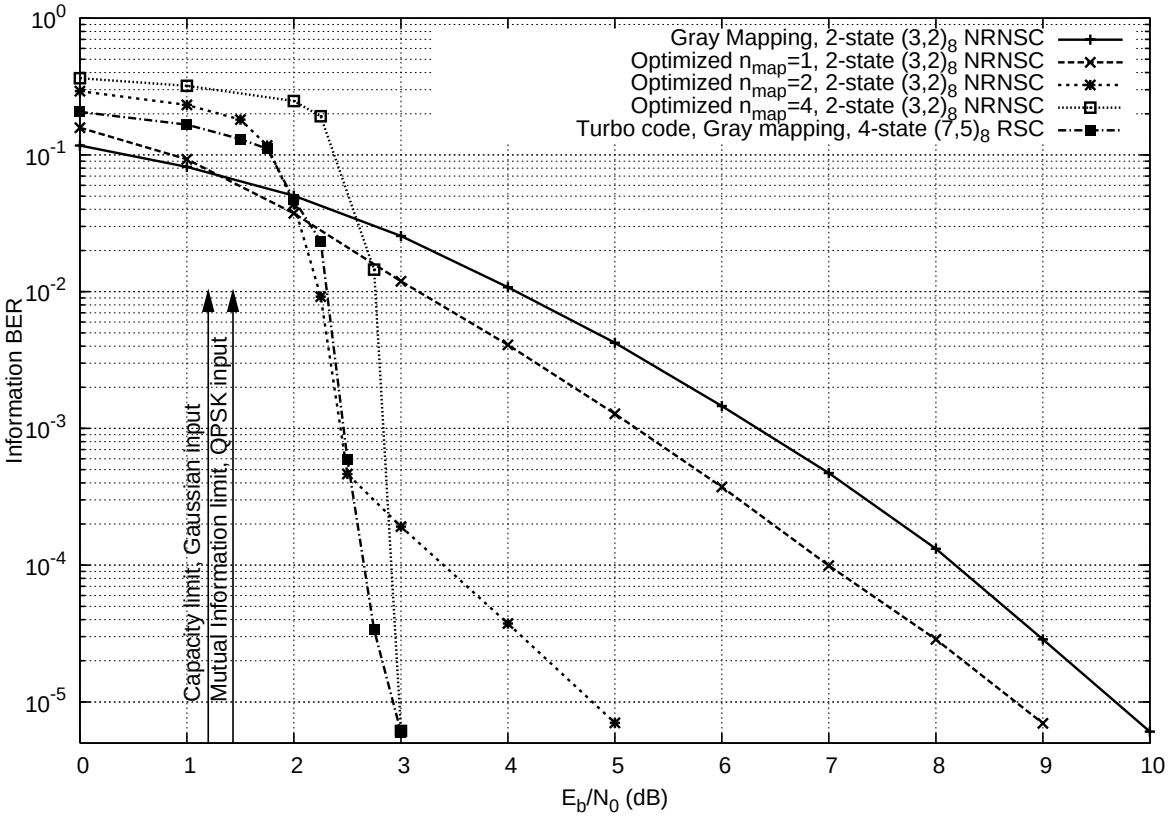

Fig. 8. Ergodic 4×4 MIMO channel, interleaver size is 8192 bits, rate-1/2 NRNSC and turbo code, QPSK modulation, 20 decoding iterations.

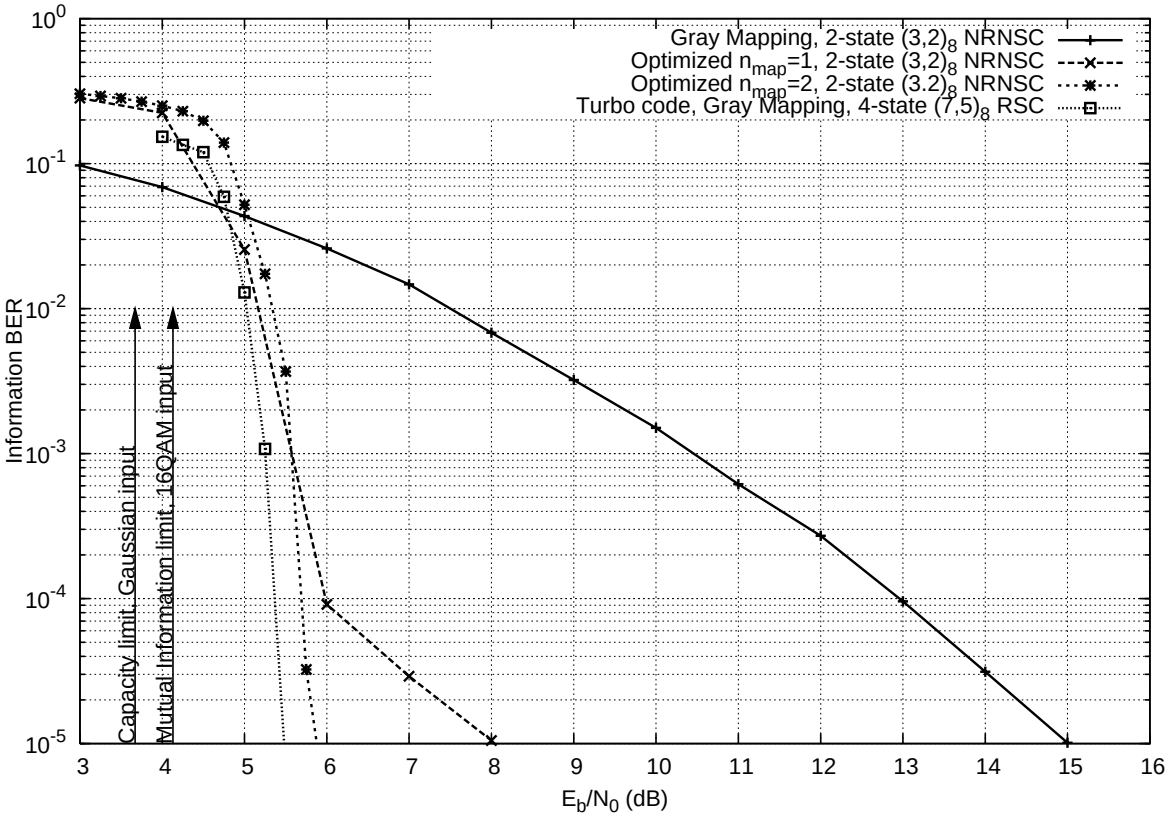

Fig. 9. Ergodic 2×2 MIMO channel, interleaver size is 8192 bits, rate-1/2 NRNSC and turbo code, 16-QAM modulation, 20 decoding iterations.