Bit-Interleaved Time-Space Codes for Fading Channels

Ezio Biglieri Giorgio Taricco Emanuele Viterbo[1]
Dipartimento di Elettronica, Politecnico
Corso Duca degli Abruzzi 24
I-10129 Torino (Italy)
e-mail: <name>@polito.it

Abstract — **Bit-interleaved coded modulation (BICM) is known to provide a robust solution to the coding problem for wireless channels, as it provides at the same time a large Hamming distance and a large (albeit nonoptimum) Euclidean distance. As BICM splits the code design into the selection of the encoder and the selection of a modulation scheme, the design task is simpler than with "standard" space-time codes.**

In this paper we investigate the performance of BICM over a multiple-antenna channel by evaluating its information-theoretic limits. Further, we provide some design guidelines and we check our results via some simulations.

I. Introduction

Recently, a sizeable literature has been developing on the ultimate performance limits of multiple-antenna systems in a fading environment (see, e.g., [7, 8, 20]). Several information-theoretic analyses have been produced: for example, [12] contains an analysis of practical transmission schemes with multiple antennas, while [20] determines the ergodic capacity of a multiple-antenna channel with independent Rayleigh fading, and [4] studies the impact of power control on system performance when channel-state information is available at the transmitter.

Given that considerable gains are achievable by a multi-antenna system, the challenge is to design coding schemes that perform close to capacity: space-time trellis codes, space-time block codes, and layered space-time codes have been advocated (see, e.g., [9, 13, 14, 15, 16, 18]).

For single-antenna sytems, a coding scheme, called *bit-interleaved coded modulation* (BICM) [3] was shown to be robust to the channel model, as it provides good performance over the Rayleigh fading channel as well as over the additive white Gaussian noise channel: in fact, it yields a large Hamming distance while at the same time approaching the capacity of high-SNR AWGN channels. As BICM separates the selection of an encoder from the selection of a modulation scheme, the design task is simpler than with "standard" space-time codes. Thus, it seems sensible to investigate how BICM performs over a multiple-antenna channel.

Since the channel model usually assumed for multiple-antenna transmission is one with block fading [1], whereby the channel gains across all antenna pairs are constant over the duration of an entire codeword, it may seem at first that BICM cannot buy much. In fact, it operates by transforming a channel into parallel *binary* channels, and it is known that over block-fading channels binary transmission is less efficient. On the other hand, recent results [11, 15] show that BICM could perform considerably well in several contexts, and hence that BICM warrants some analysis aimed at assessing its limiting performance in a multiple-antenna environment. The goal of this paper is to provide such analysis.

With our model of BICM, the output of an encoder (hereafter called the "mother encoder") is cyclically connected to t transmitters, each including a bit interleaver (whose depth is assumed infinite) followed by a modulator and an antenna. The codes seen by the transmit antennas are punctured versions of the mother code. At the receiving front-end of the channel, the bit metrics are computed, restored to their original order by a de-interleaver, and used for decoding. The receiver is assumed to have perfect knowledge of the "channel-state information," i.e., of the values taken by the channel gains.

For fairness, we compare single- and multiple-antenna schemes based on the transmission of the same number of bits per channel use, where a channel use when t transmit antennas are available consists of transmitting simultaneously from t signal constellations. Suppose for example that, irrespective of t, we want to transmit 1 bit per channel use, and that QPSK is the modulation scheme. With $t = 1$ we need a single baseline code with rate $1/2$, while with t antennas the signal transmitted by each antenna should carry $1/t$ information bits: this is obtained by using a mother code with rate $1/2t$. In general, we should compare a BICM scheme with a single transmit antenna and a rate-R baseline code against a t-transmit-antenna BICM scheme with a rate-R/t mother code.

II. Channel model

Here we consider a transmission system with t transmit and r receive antennas. Denote by L the number of symbol intervals occupied by a code word (this is called a "frame" in [16]). We write the signal received at the channel output as

$$\mathbf{y}[\ell] = \mathbf{H}[\ell]\mathbf{x}[\ell] + \mathbf{n}[\ell], \qquad \ell = 1, \ldots, L \tag{1}$$

where $\mathbf{y}[\ell]$, the signal received at discrete time ℓ, and $\mathbf{n}[\ell]$, the white noise affecting it, are $r \times 1$ vectors; $\mathbf{x}[\ell]$ denotes the $t \times 1$ vector of signals transmitted at time ℓ, selected from constellation $\mathcal{X}$; and $\mathbf{H}[\ell]$ denotes the $r \times t$ channel-gain matrix.

Further, we denote by $\mathbf{X}$ the coded sequence $(\mathbf{x}[1], \ldots, \mathbf{x}[L])$, and we write

$$\mathbf{H}[\ell] = \begin{bmatrix} \mathbf{h}_1[\ell] \\ \mathbf{h}_2[\ell] \\ \vdots \\ \mathbf{h}_r[\ell] \end{bmatrix} \tag{2}$$

where $\mathbf{h}_i[\ell]$ are $1 \times t$ vectors. Finally, we define $\mathcal{X}_i$, $i = 1, \ldots, r$, as the constellations whose signals are all possible values of $\mathbf{h}_i[\ell]\mathbf{x}[\ell]$.

The three following situations are relevant in the analysis of multiple-antenna systems:

1. $\mathbf{H}[\ell]$ is deterministic, and independent of ℓ.

2. $\mathbf{H}[\ell]$ is random, but remains constant during the transmission of a code word or of a fraction of it (the "block-fading" channel).
3. $\mathbf{H}[\ell]$ is random and takes on independent values for each ℓ (the "independent-fading" channel).

In addition, the statistics of the elements of the random channel-gain matrix determine a taxonomy of multiple-antenna channels as illustrated in the following, where we drop the index ℓ for notational simplicity:

1. $\mathbf{H}$ has independent entries $h_{ij} \sim \mathcal{N}(0, 1/2)$: this model (a channel with independent Rayleigh fading channel gains) represents a situation where both transmit and the receive antennas have sufficient spatial separation.
2. $\mathbf{H}$ has column vectors $\mathbf{h}_j^c \sim \mathcal{N}(0, \mathbf{\Sigma})$ with $\mathbf{\Sigma} = \mathbb{E}[\mathbf{h}^{c\,\dagger}\mathbf{h}^c]$: this models a channel with correlated Rayleigh fading due, for example, to insufficient separation between the transmit antennas. In expanded form, we have
$$\mathbb{E}\left[h_{ij}h_{kl}\right] = \mathbf{\Sigma}_{jl}.$$
3. $\mathbf{H}$ has row vectors $\mathbf{h}_i \sim \mathcal{N}(0, \mathbf{\Xi})$ with $\mathbf{\Xi} = \mathbb{E}[\mathbf{h}_i \mathbf{h}_i^\dagger]$: this models a channel with correlated Rayleigh fading due, for example, to insufficient separation between the receive antennas.
4. If there is joint correlation in both transmission and reception we may have $\mathrm{vec}(\mathbf{H}) \sim \mathcal{N}(0, \mathbf{\Xi}\otimes\mathbf{\Sigma})$ where $\mathbf{\Xi}\otimes\mathbf{\Sigma}$ is the Kronecker matrix product of the two correlation matrices and $\mathrm{vec}(\mathbf{H})$ is the vector of length rt obtained by stacking the columns of $\mathbf{H}$.
5. Rician fading results when there is a direct path between transmit and receive antennas. This can be modelled by a constant matrix $\mathbf{M} = \mathbb{E}[\mathbf{H}] \neq \mathbf{0}$, and results in non-central normal entries of $\mathbf{H}$.

Distinguishing between the independent-fading and the block-fading model implies the distinction between fast and slow fading. When the fading is fast, we may assume that the channel is ergodic, and hence its capacity is defined as

$$C = \mathbb{E}[\sup_{p(\mathbf{x})} \mathcal{I}(\mathbf{x};\mathbf{y} \mid \mathbf{H})] = \sup_{p(\mathbf{x})} \mathbb{E}[\mathcal{I}(\mathbf{x};\mathbf{y} \mid \mathbf{H})], \quad \mathbb{E}[\mathbf{x}^\dagger\mathbf{x}] \leq \mathcal{P} \tag{3}$$

where $\mathcal{I}(\mathbf{x};\mathbf{y} \mid \mathbf{H})$ is the *conditional mutual information* between input and output, $p(\mathbf{x})$ is the probability density of the input, and $\mathcal{P}$ is the maximum power across all transmit antennas.

When the channel is ergodic the conditional mutual information is a random variable, and the average is taken over all the channel realizations. When the fading is slow ('quasi-static' channel) the channel is nonergodic [1], and the sensible parameter to be used to assess the coding-channel quality is its *outage probability*

$$P_{\mathrm{out}}(R) = \mathbb{P}[\sup_{p(\mathbf{x})} \mathcal{I}(\mathbf{x};\mathbf{y}|\mathbf{H}) < R] \tag{4}$$

where R is the transmission rate. This represents the probability that the maximum conditional mutual information falls below R due to channel variations. In the presence of an outage, transmission is totally unreliable irrespective of the code used. We may also define the *ε-capacity* of the channel, as the rate R^* for which $P_{\mathrm{out}}(R) < \varepsilon$.

It can be shown [20] that an optimal distribution for the input $\mathbf{x}$ is in all cases $\mathcal{N}(0, \mathbf{Q})$ where $\mathbf{Q}$ is a diagonal $t \times t$ covariance matrix whose entries represent the power transmitted by each antenna (the power constraint then reduces to $\mathrm{tr}(\mathbf{Q}) < \mathcal{P}$.)

When the transmitter has the availability of channel state information, it is also possible to optimize the allocation of power so as to maximize the conditional mutual information (see [4]). We shall not consider this case here, but rather restrict our attention to a uniform power $\gamma \triangleq \mathcal{P}/t$ allocated to each transmit antenna.

III. Capacity calculations

For reference purposes, we observe first the channel capacity with Gaussian inputs: we have [20]

$$\mathcal{I}(\mathbf{x};\mathbf{y} \mid \mathbf{H}) = \log\det(\mathbf{I}_r + (\mathcal{P}/t)\mathbf{H}\mathbf{H}^\dagger) = \log\det(\mathbf{I}_t + (\mathcal{P}/t)\mathbf{H}^\dagger\mathbf{H}) \tag{5}$$

Here we are interested in discrete input signal sets $\mathcal{X}$ (for example, PSK) with $|\mathcal{X}| = 2^m$, where we assume a uniform distribution of the symbols transmitted by each antenna. Following the results and notations in [3] we have in general, with standard coded modulation (CM)

$$C = mt - \mathbb{E}_{\mathbf{x},\mathbf{y},\mathbf{H}}\left[\log_2 \frac{\sum_{\mathbf{z}\in\mathcal{X}^t} p(\mathbf{y}|\mathbf{z},\mathbf{H})}{p(\mathbf{y}|\mathbf{x},\mathbf{H})}\right] \tag{6}$$

while with BICM

$$C = mt - \sum_{i=1}^{mt} \mathbb{E}_{\mathbf{x},\mathbf{y},\mathbf{H},b}\left[\log_2 \frac{\sum_{\mathbf{z}\in\mathcal{X}^t} p(\mathbf{y}|\mathbf{z},\mathbf{H})}{\sum_{\mathbf{z}\in\mathcal{X}^t_{i,b}} p(\mathbf{y}|\mathbf{z},\mathbf{H})}\right] \tag{7}$$

where $\mathcal{X}^t$ denotes the t-fold Cartesian product of $\mathcal{X}$, and $\mathcal{X}^t_{i,b}$ denotes the subset of signals from $\mathcal{X}^t$ whose ith bit is equal to b. We also have, under our assumptions, the channel transition probabilities

$$p(\mathbf{y} \mid \mathbf{x},\mathbf{H}) = (\pi N_0)^{-t}\exp(-\|\mathbf{y} - \mathbf{H}\mathbf{x}\|^2/N_0)$$

Numerical results. Fig. 1 shows the cannel capacity with a QPSK constellation in conjunction with CM and BICM (Gray mapping is used [3]). The capacity for Gaussian inputs is also shown for comparison. The channel is ergodic, with fast Rayleigh fading ($r, t = 1, 2$). Notice that as the signal-to-noise ratio (SNR) increases the capacities with QPSK tend to $2t$ bits per channel use, as t signals are simultaneously transmitted. If, for example, two transmit antennas are used to transmit 2 bits per channel use it is apparent that a much lower SNR is needed than with one antenna: this explains the coding gain afforded by multiple antennas. With $t = 1$ the capacities of CM and BICM are practically identical, while for $t = 2$ the performance loss of BICM is apparent. Similar considerations, *mutatis mutandis*, hold for 8-PSK, as shown in Fig. 2. Notice in particular the large discrepancy between the capacity curves with $t = 2$, $r = 1$ at intermediate SNR.

Figs. 3 to 6 show the ε-capacity of a multiple-antenna system with a QPSK constellation in conjunction with with CM and BICM and different values of ε. (Gray mapping is used [3]). The channel is non-ergodic, with Rayleigh fading ($r, t = 1, 2$).

IV. Error probabilities

In this section we provide some considerations about the design of mother codes for the BICM scheme studied previously.

By paralleling the calculations in [3, p. 936], we can obtain a general expression for the pairwise error probability $P(\mathbf{C} \to \widehat{\mathbf{C}})$, where $\mathbf{C}$, $\widehat{\mathbf{C}}$ denote two code words. This is asymptotically upperbounded (for high SNR) as follows:

$$P(\mathbf{C} \to \widehat{\mathbf{C}}) \overset{\sim}{\leq} \nu \max_{\mathbf{X},\widehat{\mathbf{X}}} P(\mathbf{X} \to \widehat{\mathbf{X}}) \tag{8}$$

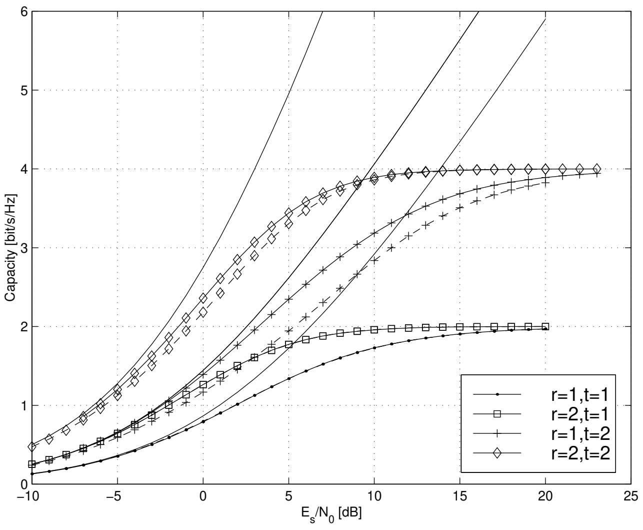

Figure 1: Capacity of a t-transmit, r-receive antenna system CM (continuous lines) and BICM (dashed lines) for a QPSK constellation transmitted over the ergodic Rayleigh fading channel. The capacity with Gaussian inputs is also shown.

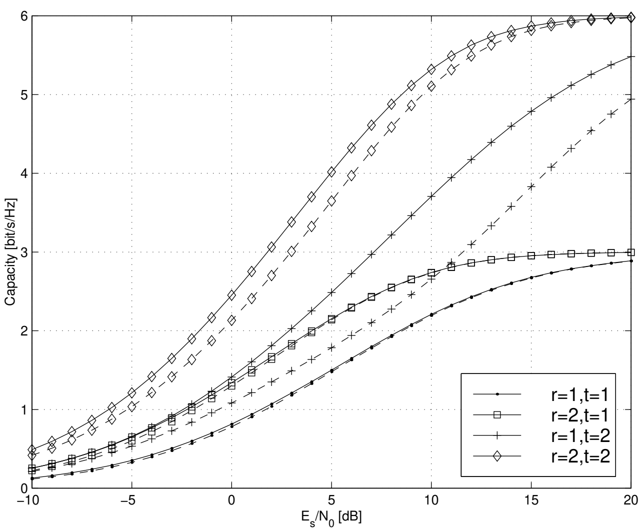

Figure 2: Capacity of 8-PSK with CM and BICM for the ergodic Rayleigh fading channel

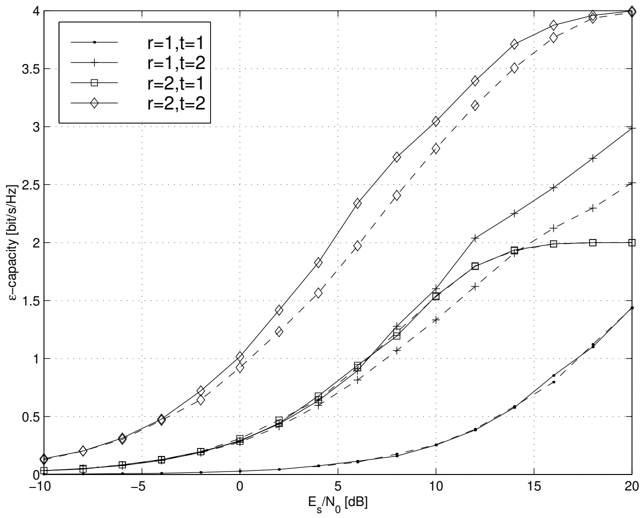

Figure 3: ε-capacity CM and BICM with QPSK and with Gaussian inputs for the non-ergodic Rayleigh fading channel ($r, t = 1, 2$) at given outage probability $\varepsilon = 0.02$.

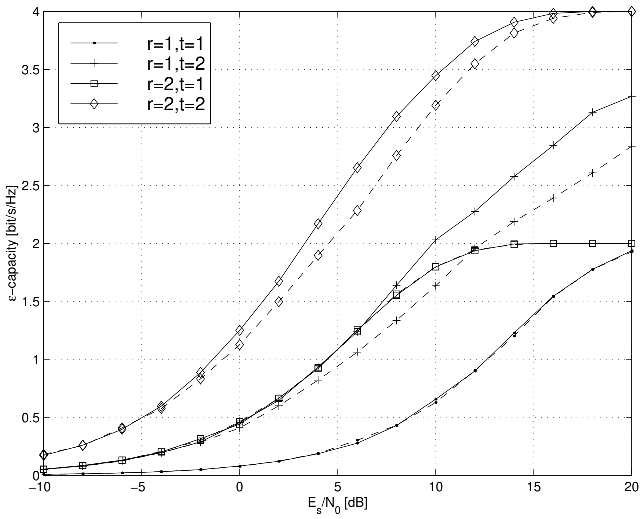

Figure 4: ε-capacity of QPSK with CM, BICM, and Gaussian input for the non-ergodic Rayleigh fading channel ($r, t = 1, 2$) at given outage probability $\varepsilon = 0.05$.

where

$$P(\mathbf{X} \to \widehat{\mathbf{X}}) \triangleq \mathbb{E}_{\mathbf{x},\mathbf{H}} \left[P(\mathbf{X} \to \widehat{\mathbf{X}} \mid \mathbf{H}) \right] \tag{9}$$

and $\mathbf{X}$, $\widehat{\mathbf{X}}$ denote the possible signal sequences compatible with the code words $\mathbf{C}$, $\widehat{\mathbf{C}}$, respectively, and with the labeling maps, and ν is a constant.

Deterministic channel. With a deterministic channel each receive antenna observes a signal which is a linear combination, with known coefficients (the channel-state information), of the signals transmitted by t antennas. This is tantamount to saying that the signals observed by the receive antennas are chosen from the constellations $\mathcal{X}_i$.

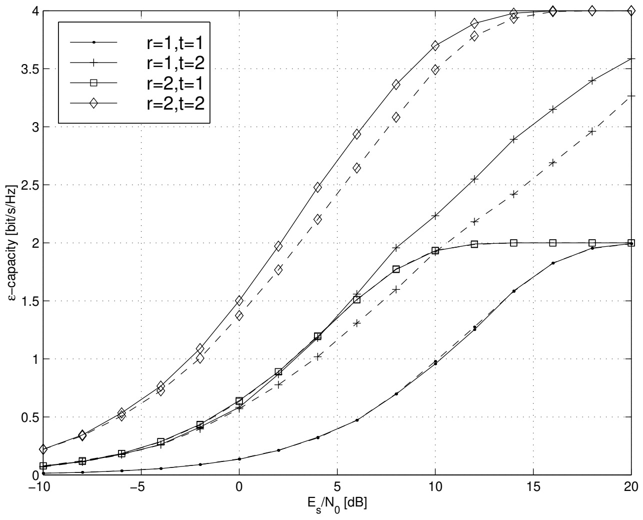

Figure 5: ε-capacity of QPSK with CM, BICM, and Gaussian input for the non-ergodic Rayleigh fading channel $(r, t = 1, 2)$ at given outage probability $\varepsilon = 0.1$.

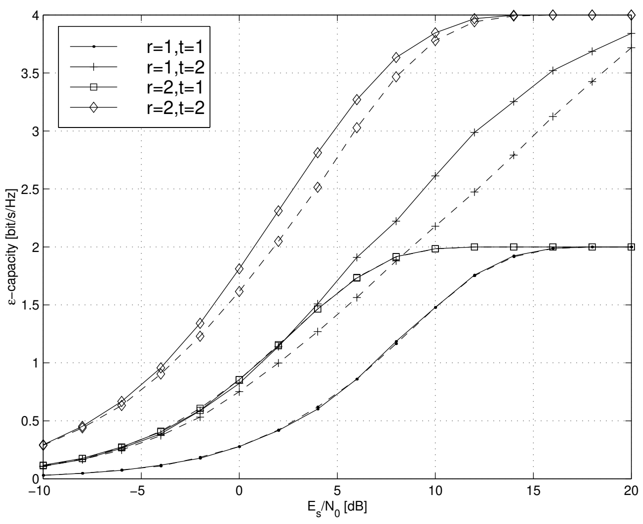

Figure 6: ε-capacity of QPSK with CM, BICM, and Gaussian input for the non-ergodic Rayleigh fading channel $(r, t = 1, 2)$ at given outage probability $\varepsilon = 0.2$.

The Bhattacharyya bound yields in this case

$$P(\mathbf{X} \to \widehat{\mathbf{X}} \mid \mathbf{H}) \leq \exp\left(-\delta^2(\mathbf{X},\, \widehat{\mathbf{X}})/4N_0\right) \qquad (10)$$

where $\delta^2(\mathbf{X},\, \widehat{\mathbf{X}})$, the Euclidean distance between $\mathbf{X}$ and $\widehat{\mathbf{X}}$, is given by

$$\begin{aligned} \delta^2(\mathbf{X}, \widehat{\mathbf{X}}) &= \sum_{\ell=1}^{L}\sum_{j=1}^{r} |\mathbf{h}_j[\ell](\mathbf{x}[\ell] - \widehat{\mathbf{x}}[\ell])|^2 \\ &= \sum_{j=1}^{r}\sum_{\ell=1}^{L} \mathbf{h}_j[\ell](\mathbf{x}[\ell] - \widehat{\mathbf{x}}[\ell])(\mathbf{x}[\ell] - \widehat{\mathbf{x}}[\ell])^\dagger \mathbf{h}_j[\ell]^\dagger \end{aligned} \qquad (11)$$

Now, since $\mathbf{H}[\ell]$ is constant here, we can write

$$\delta^2(\mathbf{X}, \widehat{\mathbf{X}}) = \sum_{j=1}^{r} \mathbf{h}_j \mathbf{A} \mathbf{h}_j^\dagger \qquad (12)$$

where

$$\mathbf{A} \triangleq \sum_{\ell=1}^{L} (\mathbf{x}[\ell] - \widehat{\mathbf{x}}[\ell])(\mathbf{x}[\ell] - \widehat{\mathbf{x}}[\ell])^\dagger \qquad (13)$$

(the dagger † denotes Hermitian transpose). Since $\mathbf{A}$ is a $t \times t$ nonnegative definite matrix, we can write it as $\mathbf{A} = \mathbf{Q}\mathbf{D}\mathbf{Q}^\dagger$ where $\mathbf{Q}$ is a unitary matrix and $\mathbf{D}$ is the diagonal matrix of the (nonnegative) eigenvalues of $\mathbf{A}$. Thus, we have

$$\begin{aligned} \delta^2(\mathbf{X}, \widehat{\mathbf{X}}) &= \sum_{j=1}^{r} (\mathbf{h}_j\mathbf{Q})\mathbf{D}(\mathbf{h}_j\mathbf{Q})^\dagger = \sum_{j=1}^{r}\sum_{i=1}^{t} \beta_{ji}\lambda_i(\mathbf{A}) \\ &\geq \min_{1\leq i \leq t, \mathbf{A}} \sum_{j=1}^{r} \beta_{ji} \operatorname{Tr}(\mathbf{A}) \end{aligned} \qquad (14)$$

where β_{ji} is the squared magnitude of the ith component of the vector $\mathbf{h}_j\mathbf{Q}$ and $\lambda_i(\mathbf{A})$ is the ith eigenvalue of $\mathbf{A}$ $(i = 1, \ldots, t)$. Moreover, we have

$$\begin{aligned} \operatorname{Tr}(\mathbf{A}) &= \sum_{\ell=1}^{L} |\mathbf{x}[\ell] - \widehat{\mathbf{x}}[\ell]|^2 = \sum_{\ell=1}^{L}\sum_{i=1}^{t} |x_i[\ell] - \widehat{x}_i[\ell]|^2 \\ &\geq \sum_{i=1}^{t} d_{i,\min}\delta^2(\mathcal{X}) \end{aligned} \qquad (15)$$

where $d_{i,\min} \triangleq \min_{\mathbf{X}\neq\widehat{\mathbf{X}}} \sum_{\ell=1}^{L} \mathbf{1}\{x_i[\ell] \neq \widehat{x}_i[\ell]\}$ is the minimum Hamming distance between the codewords output from the ith antenna and $\delta^2(\mathcal{X}) \triangleq \min_{x,\widehat{x}\in\mathcal{X}, x\neq\widehat{x}} |x - \widehat{x}|^2$ is the minimum Euclidean distance of the constellation $\mathcal{X}$. Finally, we obtain the lower bound

$$\delta^2(\mathbf{X}, \widehat{\mathbf{X}}) \geq \left\{ \min_{1\leq i\leq t, \mathbf{A}} \sum_{j=1}^{r} \beta_{ji} \right\} \sum_{i=1}^{t} d_{i,\min}\delta^2(\mathcal{X}) \qquad (16)$$

Unfortunately, $\min_{1\leq i\leq t,\mathbf{A}} \sum_{j=1}^{r} \beta_{ji}$ is difficult to calculate because it depends on the matrix $\mathbf{A}$ and may even be zero, thus turning the bound to a trivial one. In order for this to occur, the matrix $\mathbf{HQ}$ must exhibit an all-zero column. Hence, the linear system $\mathbf{H}\mathbf{q}_i = \mathbf{0}$ must have a nonzero solution because the columns of the unitary matrix $\mathbf{Q}$ cannot be $\mathbf{0}$, i.e., have all zero entries. This happens if and only if the rank of $\mathbf{H}$ is smaller than t [10, Sec. 0.4]. Otherwise, if the rank of $\mathbf{H}$ is not smaller than t, $\min_{1\leq i\leq t,\mathbf{A}} \sum_{j=1}^{r} \beta_{ji} > 0$.

From the above we can figure out the contribution of the various parameters influencing the performance of multiple-antenna BICM:

- The error exponent increases at least proportionally to $\delta^2(\mathcal{X})$: this is the minimum diversity achieved by increasing the number of receiving antennas.

- By increasing the number t of transmit antennas the error exponent increases at least as the sum of the Hamming distances of the punctured versions of the mother code. Thus, diversity gain due to multiple transmit antenna is achieved to the extent that this sum exceeds that of the baseline code (whose rate is t times that of the mother code).

Block-fading channel. Again, $\mathbf{H}$ does not depend on ℓ, and we have from (11):

$$\delta^2(\mathbf{X},\, \widehat{\mathbf{X}}) = \sum_{j=1}^{r} \mathbf{h}_j \mathbf{A} \mathbf{h}_j^\dagger \tag{17}$$

where $\mathbf{A}$ is as in (13).

$$\mathbf{A} \triangleq \sum_{\ell=1}^{L} (\mathbf{x}[\ell] - \hat{\mathbf{x}}[\ell])(\mathbf{x}[\ell] - \hat{\mathbf{x}}[\ell])^\dagger \tag{18}$$

Following the footsteps of Tarokh *et al.* [16], and assuming that the entries of $\mathbf{H}$ are independent, circularly-distributed zero-mean complex Gaussian RVs with unit variance (i.e., $\mathbb{E}|h_{ji}|^2 = 1$) we have

$$\begin{aligned} P(\mathbf{X} \to \widehat{\mathbf{X}}) &\triangleq \mathbb{E}_{\mathbf{x},\mathbf{H}}[P(\mathbf{X} \to \hat{\mathbf{X}} \mid \mathbf{H})] \\ &\le \left[\frac{1}{\prod_{i=1}^{t}(1+\lambda_i/4N_0)} \right]^r \\ &= \det(\mathbf{I} + \mathbf{A}/4N_0)^{-r} \end{aligned} \tag{19}$$

where λ_i are the eigenvalues of $\mathbf{A}$.

We upperbound (19) by using the following bound, especially useful for low SNR and valid for every $n \times n$ nonnegative-definite matrix $\mathbf{A}$:

$$\det(\mathbf{I} + \mathbf{A}) = \prod_{i=1}^{n}(1+\lambda_i(\mathbf{A})) \ge 1 + \sum_{i=1}^{n} \lambda_i(\mathbf{A}) = 1 + \mathrm{tr}(\mathbf{A})$$

where $\lambda_i(\cdot)$ denotes the ith eigenvalue and $\mathrm{tr}(\cdot)$ denotes the trace of a matrix.

Application of this bound yields

$$\begin{aligned} P(\mathbf{X} \to \widehat{\mathbf{X}}) &\le [1 + \mathrm{tr}(\mathbf{A})/4N_0]^{-r} \\ &= \left[1 + \frac{1}{4N_0} \sum_{\ell=1}^{L} |\mathbf{x}[\ell] - \hat{\mathbf{x}}[\ell]|^2 \right]^{-r} \\ &\le \left[1 + \frac{\delta^2(\mathcal{X})}{4N_0} \sum_{i=1}^{t} d_{i,\min} \right]^{-r} \end{aligned}$$

In this case:

- The error exponent increases proportionally to r: thus, a diversity equal to the number of receive antennas is always achieved.
- By increasing the number t of transmit antennas we obtain at least a coding gain proportional to the sum of the minimum Hamming distances of the punctured versions of the mother code.

Independent-fading channel As with the block-fading channel, we assume that the matrices $\mathbf{H}[\ell]$ are iid with independent entries. Following [16] we have

$$\Pr(\mathbf{X} \to \widehat{\mathbf{X}}) \le \mathbb{E}_{\mathbf{x}} \left[\exp\left(-\frac{1}{4N_0} \sum_{\ell=1}^{L} \sum_{j=1}^{r} \mathbf{h}_j[\ell] \mathbf{A}[\ell] \mathbf{h}_j[\ell]^\dagger \right) \right] \tag{20}$$

where

$$\mathbf{A}[\ell] \triangleq (\mathbf{x}[\ell] - \widehat{\mathbf{x}}[\ell])(\mathbf{x}[\ell] - \widehat{\mathbf{x}}[\ell])^\dagger \tag{21}$$

Noting that $\mathbf{A}[\ell]$ has only one nonzero eigenvalue, and this is equal to $|\mathbf{x}[\ell] - \widehat{\mathbf{x}}[\ell]|^2$, we obtain, after diagonalization of $\mathbf{A}[\ell]$:

$$\begin{aligned} \Pr(\mathbf{X} \to \widehat{\mathbf{X}}) &\le \prod_{\ell=1}^{L} \left(1 + \frac{|\mathbf{x}[\ell] - \widehat{\mathbf{x}}[\ell]|^2}{4N_0} \right)^{-r} \\ &\le \left(1 + \left\{ \min_{1 \le i \le t} d_{i,\min} \right\} \frac{\delta^2(\mathcal{X})}{4N_0} \right)^{-r d_{\min}} \end{aligned} \tag{22}$$

where $d_{\min} \triangleq \min_{\mathbf{X},\widehat{\mathbf{X}}} \sum_{\ell=1}^{L} \mathbf{1}\{\mathbf{x}[\ell] \neq \widehat{\mathbf{x}}[\ell]\}$ is the minimum Hamming distance of the code interpreted as a collection of vectors of length t.

A comparison. For large SNR (i.e., when $N_0 \to 0$) the code diversity achieved with the block-fading channel is equal to θr (where θ is the rank of the matrix $\mathbf{A}$) while the code diversity achieved with the independent fading channel is equal to $r d_{\min}$ (where $d_{\min}$ is the minimum Hamming distance of the code interpreted as a collection of vectors of length t).

While the value of θ can never exceed t, the size of $\mathbf{A}$, $d_{\min}$ is only limited by the complexity of the code. Thus, as expected, we usually have a better performance over the independent fading channel than over the block-fading channel.

V. Numerical results

The results above seem to indicate that a sensible parameter to be taken into consideration in the selection of BICM schemes is the sum of the Euclidean distances of the codes obtained as projections of a mother code over t transmit branches. However, since this indication was based on bounds (that might be loose at times), additional effort is required to validate it.

Here we show one simulation example, based on transmission over the block-fading channel. We consider frames of length 130 symbols carrying a total of 130 information bits each. As a baseline we choose a single-transmit-, single-receive-antenna system with 4-PSK and a rate-1/2 binary convolutional code with octal generators $(15, 17)$, 8 states, and maximum free distance ($d_{\text{free}} = 6$). This yields a spectral efficiency of 1 bit per channel use. For BICM, the system has $t = 2$, $r = 1$, and 4-PSK. We choose three different rate-1/4 binary convolutional codes, denoted C_a, C_b, and C_c, with octal generators $(13, 13, 15, 17)$, $(15, 13, 15, 17)$, and $(15, 17, 15, 17)$, respectively. Such a system offers the same number of bits per channel use as the baseline. Note that C_c is equivalent to having two identical convolutional codes with rate $1/2$ and generators $(15, 17)$ on each TX antenna. Consideration of C_c shows that independent coding on each TX antenna is not an effective solution.

Figure 7 shows frame error probabilities, from which we observe the following:

- All $t = 2$ curves have the same slope, which is greater than for $t = 1$ due to the transmit diversity;

- Both C_a and C_b outperform the baseline system showing gains of 5 dB and 3.5 dB, respectively, for a frame error probability 10^{-3};
- Code C_c is the least effective.

Similar overall performance at a lower E_b/N_0 was observed when $r = 2$ was chosen in both baseline and BICM systems.

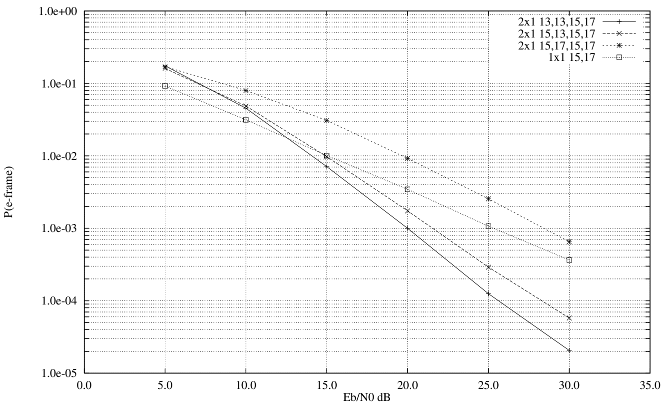

Figure 7: Frame error probability over the slow fading channel

VI. Conclusions

We have investigated the performance of BICM over a multiple-antenna channel by evaluating its information-theoretic limits. Design guidelines, to be validated by simulation, were also indicated.

References

[1] E. Biglieri, J. Proakis, and S. Shamai (Shitz), "Fading channels: Information-theoretic and communications aspects," *IEEE Trans. Inform. Theory*, Vol. 44, No. 6, pp. 2619–2692, October 1998.

[2] F. Boixadera Espax and J. J. Boutros, "Capacity considerations for wireless multiple-input multiple-output channels," *1999 Workshop on Multiaccess, Mobility and Teletraffic for Wireless Communications (MMT'99)*, Venezia, Italy, October 6–8, 1999.

[3] G. Caire, G. Taricco, and E. Biglieri, "Bit-interleaved coded modulation," *IEEE Trans. Inform. Theory*, Vol. 44, pp. 927–946, May 1998.

[4] G. Caire, G. Taricco, and E. Biglieri, "Limiting performance of block-fading channels with multiple antennas," *IEEE Trans. Inform. Theory*, accepted for publication, 2000.

[5] C.-N. Chuah, D. Tse, and J. M. Kahn, "Capacity of multi-antenna array systems in indoor wireless environment," *GLOBECOM'98*, Sidney, Australia, Nov. 8–12, 1998.

[6] G. J. Foschini, "Layered space-time architecture for wireless communication in a fading environment when using multi-element antennas," *Bell Labs Tech. J.*, Vol. 1, No. 2, pp. 41–59, Autumn 1996.

[7] G. J. Foschini and M. J. Gans, "On limits of wireless communications in a fading environment when using multiple antennas," *Wireless Personal Communications*, Vol. 6, No. 3, pp. 311–335, March 1998.

[8] G. J. Foschini and R. A. Valenzuela, "Initial estimation of communication efficiency of indoor wireless channels," *Wireless Networks*, Vol. 3, No. 2, pp. 141–154, 1997.

[9] J.-C. Guey, M. P. Fitz, M. R. Bell, and W.-Y. Kuo, "Signal design for transmitter diversity wireless communication systems over Rayleigh fading channels," *IEEE Trans. Commun.*, Vol. 47, No. 4, pp. 527–537, April 1999.

[10] R. Horn and C. Johnson, *Matrix Analysis*. New York: Cambridge University Press, 1985.

[11] C. Lamy, F. Boixadera, and J. Boutros, "Iterative APP decoding and channel estimation for multiple-input multiple-output channels," January 2000, *to be published*.

[12] A. Narula, M. D. Trott, and G. Wornell, "Performance limits of coded diversity methods for transmitter antenna arrays," *IEEE Trans. on Inform. Theory,* Vol. 45, No. 7, pp. 2418–2433, Nov. 1999.

[13] D. Shiu and J. M. Kahn, "Design of high-throughput codes for multiple-antenna wireless systems," Submitted to *IEEE Trans. Inform. Theory*, January 1999.

[14] D. Shiu and J. M. Kahn, "Layered space-time codes for wireless communications using multiple transmit antennas," *IEEE International Conf. Telecomm. (ICC'99)*, Vancouver, BC, June 6–10, 1999.

[15] A. S. Stefanov and T. M. Duman, "Turbo coded modulation for systems with transmit and receive antenna diversity," *Proc. 1999 Global Telecomm. Conf. (GLOBECOM'99)*, Rio de Janeiro, Brazil, pp. 2336–2340, Dec. 5–9, 1999.

[16] V. Tarokh, N. Seshadri, and A. R. Calderbank, "Space-time codes for high data rate wireless communication: Performance criterion and code construction," *IEEE Trans. on Inform. Theory* Vol. 44, No. 2, pp. 744 – 765, March 1998.

[17] V. Tarokh, A. Naguib, N. Seshadri, and A. R. Calderbank, "Space-time codes for high data rates wireless communication: Performnace criteria in the presence of channel estimation errors, mobility, and multiple paths," *IEEE Trans. Commun.*, Vol. 47, No. 2, pp. 199–207, Feb. 1999.

[18] V. Tarokh, H. Jafarkhani, and A. R. Calderbank, "Space-time block codes from orthogonal designs," *IEEE Trans. Inform. Theory*, Vol. 45, No. 5, pp. 1456–1467, July 1999.

[19] V. Tarokh, A. Naguib, N. Seshadri, and A. R. Calderbank, "Combined array processing and space-time coding," *IEEE Trans. Inform. Theory*, Vol. 45, No. 4, pp. 1121–1128, May 1999.

[20] İ. E. Telatar, "Capacity of multi-antenna Gaussian channels," *European Trans. Telecomm.*, Vol. 10, No. 6, pp. 585–595, November–December 1999.

[21] A. Yongacoglu and M. Siala, "Space-time codes for fading channels," *Proc. VTC-1999 Fall*, Amsterdam, The Netherlands, pp. 2495–2499, Sept. 18–22, 1999.