International Symposium on Information Theory and Its Applications
Honolulu, Hawaii, U.S.A., November 5–8, 2000

Information-Theoretic Analysis of Bit-Interleaved Time-Space Codes for Fading Channels

Ezio Biglieri Giorgio Taricco Emanuele Viterbo

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.

1. Introduction

Recently, a sizeable literature has been developing on the ultimate performance limits of multiple-antenna systems in a fading environment (see, e.g., [4, 5, 14]). Several information-theoretic analyses have been produced: for example, [8] contains an analysis of practical transmission schemes with multiple antennas, while [14] determines the ergodic capacity of a multiple-antenna channel with independent Rayleigh fading, and [3] studies the impact of power control on system performance when channel-state information is also 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., [6, 9, 10, 11, 12, 13]).

For single-antenna sytems, a coding scheme, called *bit-interleaved coded modulation* (BICM) [2] 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 [7, 11] 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 bit/s/Hz.

2. 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 [12]). 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).

When $\mathbf{H}$ is random, it is commonly assumed that it has independent entries h_{ij}, each circularly complex Gaussian distributed with zero mean and variance $\mathbb{E}[|h_{ij}|^2] = 1$: this model (a channel with independent Rayleigh fading channel gains) represents a situation where both transmit and the receive antennas have sufficient spatial separation.

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

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

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*

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

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 maximum rate $R^*(\varepsilon)$ for which $P_{\text{out}}(R) < \varepsilon$.

It can be shown [14] 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 $\text{tr}(\mathbf{Q}) < \mathcal{P}$.)

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 [3]). 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.

3. Capacity calculations

For reference purposes, we observe first the channel capacity with Gaussian inputs: we have [14]

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

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 [2] 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

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

4. Numerical results.

Fig. 1 shows the cannel capacity with a QPSK constellation in conjunction with CM and BICM (Gray mapping is used [2]). The capacity for Gaussian inputs is also shown for comparison. The channel is ergodic, with fast Rayleigh fading $(r, t = 1, 2)$.

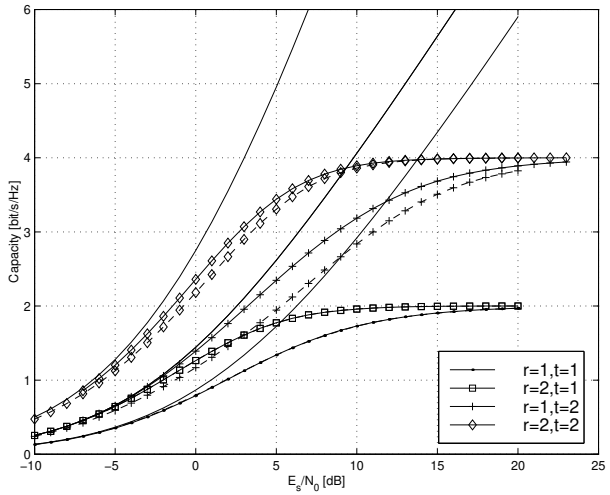

Figure 1: Capacity of a t-transmit, r-receive antenna system CM (continuous lines) and BICM (dashed lines) for a QPSK constellation transmitted over the ergodic Rayleigh fading channel. The capacity with Gaussian inputs is also shown.

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. Notice in particular the large discrepancy between the capacity curves with $t = 2$, $r = 1$ at intermediate SNR.

Figures 2 and 3 show the ε-capacity of a multiple-antenna system with a QPSK constellation in conjunction with with CM and BICM and $\varepsilon = 0.02$ and 0.05. (Gray mapping is used [2]). The channel is non-ergodic, with Rayleigh fading $(r, t = 1, 2)$.

As an example of an actual BICM design, let us consider a rate $R = 1/3$ turbo coded QPSK BICM scheme for the cases $(t = 2, r = 1)$ and $(t = 1, r = 1)$. To make a fair comparison, in both cases we consider frames with 416 information bits and a spectral efficiency $\eta = 4/3$ bit/s/Hz. A coding scheme with the above parameters is found in UMTS systems where a

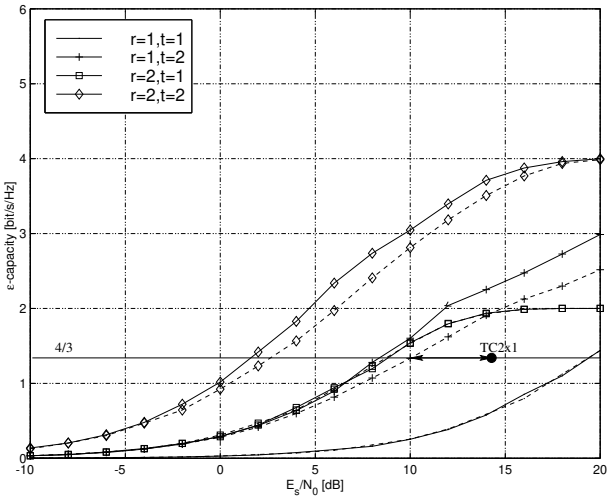

Figure 2: ε-capacity CM (continuous lines) and BICM (dashed lines) with QPSK for the non-ergodic Rayleigh fading channel $(r, t = 1, 2)$ at given outage probability $\varepsilon = 0.02$.

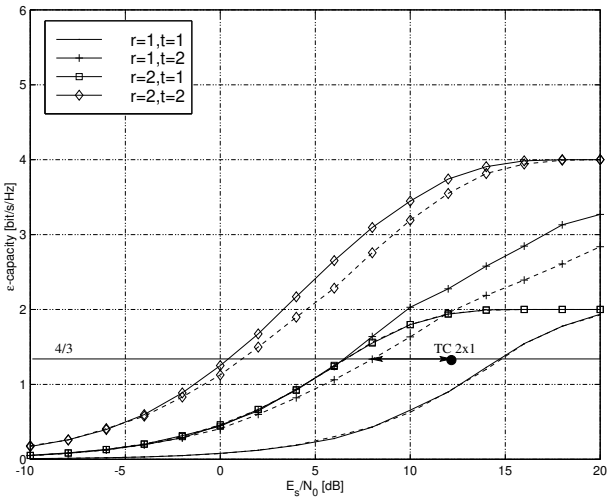

Figure 3: ε-capacity of QPSK with CM (continuous lines) and BICM (dashed lines) for the non-ergodic Rayleigh fading channel $(r, t = 1, 2)$ at given outage probability $\varepsilon = 0.05$.

single turbo codeword is mapped onto one "slot" of 1248 bits at the highest bit rate (1920 kbps), corresponding to a spreading factor 4. For $(t = 2, r = 1)$ we have $L = 312$ channel uses at rate R_s and for $(t = 1, r = 1)$ we have $L = 624$ channel uses at rate $2R_s$. Figure 4 shows the performance in terms of frame and bit error rates (FER and BER) as a function of E_b/N_0 on a slow fading channel. The diversity gain of the two transmit antenna scheme is evident from the increased slope of the error curves.

In order to compare these results with the ε-capacity

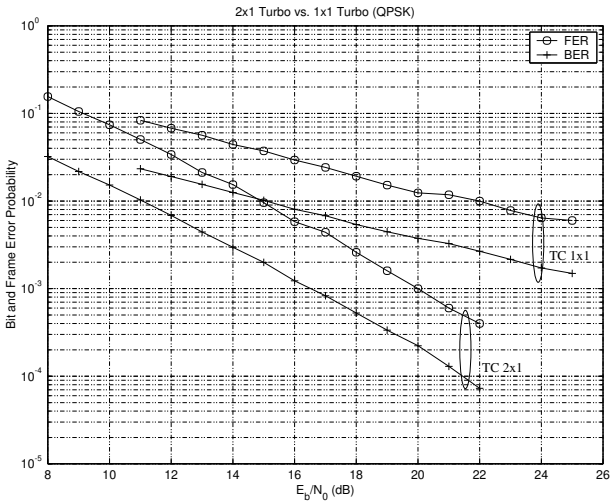

Figure 4: FER and BER performance of turbo coded rate 1/3 BICM with $t = 1, 2$ and $r = 1$

curves we find the $E_s/N_0 = \eta E_b/N_0$ values giving a FER=ε. Hence, for $t = 2, r = 1$ we have FER=0.02 at $E_s/N_0 = 14.25$ dB and FER=0.05 at $E_s/N_0 = 12.25$ dB. Figures 2 and 3 show that the turbo coded scheme (TC2x1) is about 4.25 dB away from the corresponding BICM ε-capacity curve.

Similar reasoning can be used to evaluate the performance agaist the information-theoretic limits of all practical coding schemes for multiple antenna systems presented in the literature.

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

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

[4] 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.

[5] 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.

[6] 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.

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

[8] 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.

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

[10] 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.

[11] 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.

[12] 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.

[13] 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.

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