

Impact of Signal Constellation Expansion on the Achievable Diversity of Pragmatic Bit-interleaved Space-Time Codes

Albert Guillén i Fàbregas and Giuseppe Caire

Abstract—This letter studies the effect of signal constellation expansion on the achievable diversity of pragmatic bit-interleaved space-time codes in quasistatic multiple antenna channels. Signal constellation expansion can be obtained either by increasing the size of the constellation in the complex plane or by using multidimensional linear mappings. By means of two simple constructions, we provide a comparison of the two options with message passing decoding. We show that multidimensional expansion achieves some performance advantage over complex-plane expansion at the cost of significantly higher decoding complexity and larger peak-to-average power ratio of the transmitted signals.

Index Terms—Modulation and coding, MIMO systems, iterative detection, bit-interleaved coded modulation, concatenated codes, space-time coding, fading channel, outage probability.

I. INTRODUCTION AND SYSTEM MODEL

MULTIPLE antenna transmission has emerged as a key technology to achieve large spectral and power efficiency in wireless communications. In this work, we consider communication in a multiple antenna environment with N_T transmit and N_R receive antennas in quasistatic frequency flat fading. The complex baseband received signal matrix $\mathbf{Y} \in \mathbb{C}^{N_R \times L}$ is given by,

$$\mathbf{Y} = \sqrt{\rho} \mathbf{H} \mathbf{X} + \mathbf{Z}, \quad (1)$$

where L is the block length, $\mathbf{X} = [\mathbf{x}_1 \dots \mathbf{x}_{N_T}]^T \in \mathbb{C}^{N_T \times L}$ is the transmitted signal matrix, $\mathbf{H} = [\mathbf{h}_1 \dots \mathbf{h}_{N_T}] \in \mathbb{C}^{N_R \times N_T}$, is the fading channel matrix which stays constant during the whole transmission of \mathbf{X} (quasistatic fading), $\mathbf{Z} \in \mathbb{C}^{N_R \times L}$ is a matrix of noise samples i.i.d. $\sim \mathcal{N}_{\mathbb{C}}(0, 1)$, and ρ is the average signal to noise ratio (SNR) per transmit antenna. The elements of \mathbf{H} are assumed to be i.i.d. circularly symmetric Gaussian random variables $\sim \mathcal{N}_{\mathbb{C}}(0, 1)$ (frequency flat Rayleigh fading). The channel is assumed to be perfectly known at the receiver and not known at the transmitter.

Manuscript received October 29, 2004; revised April 7, 2005; accepted June 18, 2005. The associate editor coordinating the review of this paper and approving it for publication was A. Swami. This research was supported in part by the ANTIPODE project of the French Telecommunications Research Council RNRT, and by Institut Eurecom's industrial partners: Bouygues Telecom, Fondation d'entreprise Groupe Cegetel, Fondation Hasler, France Telecom, Hitachi, STMicroelectronics, Swisscom, Texas Instruments and Thales. This work has been presented in part at the 2004 European Signal Processing Conference EUSIPCO, Vienna, Austria, September 2004.

The authors were with the Mobile Communications Department, Institut EURECOM, 2229, Route des Cretes, 06904 Sophia Antipolis Cedex, France. A. Guillén i Fàbregas is now with the Institute for Telecommunications Research, University of South Australia, SPRI Building - Mawson Lakes Blvd., Mawson Lakes SA 5095, Australia (e-mail: albert.guillen@unisa.edu.au). G. Caire is now with the Department of Electrical Engineering, University of Southern California, 3740 McClintock Ave., Los Angeles, CA 90089.

Digital Object Identifier 10.1109/TWC.2006.04734.

The multiple-input multiple-output (MIMO) channel defined by (1) has zero capacity. A Space-Time Code (STC) $\mathcal{S} \subseteq \mathbb{C}^{N_T \times L}$ is a coding scheme that exploits both temporal and space dimensions in order to achieve reliable communication. Consider an ensemble of STCs generated according to the input distribution P_X . We denote the mutual information per channel use (for a fixed \mathbf{H}) by $I_{\mathbf{H}}(P_X)$. It can be shown that the minimum achievable error probability for the ensemble in the limit for large block length is given by the *information outage probability* defined as $P_{\text{out}}(R) = \Pr(I_{\mathbf{H}}(P_X) \leq R)$ where R is the transmission rate in bits per channel use. When $P_X = \mathcal{N}_{\mathbb{C}}(0, 1)$ (Gaussian inputs), $I_{\mathbf{H}}(P_X) = \log \det(\mathbf{I} + \rho \mathbf{H}^H \mathbf{H})$. The goodness of a given STC is usually measured by its ability to approach the outage probability limit.

Conventional code design for quasistatic MIMO channels is based on the ML decoding union bound error probability [1]. The average pairwise error probability, i.e., the probability of deciding in favor of \mathbf{X}' when \mathbf{X} was transmitted assuming that there are no other codewords, for large ρ behaves as $G_c \rho^{-d_r N_R}$, where G_c is the *coding gain* and $d_r N_R$ denotes the diversity gain, where

$$d_r = \min_{\mathbf{X}, \mathbf{X}' \in \mathcal{S}} \text{rank}(\mathbf{X} - \mathbf{X}') \quad (2)$$

is the *rank diversity* of the STC $\mathcal{S} \subseteq \mathbb{C}^{N_T \times L}$. Conventional STC design searches for *full-diversity* codes \mathcal{S} , i.e., $d_r = N_T$, with the largest possible coding gain.

In this work, we study two different approaches to construct pragmatic STCs with full-diversity performance. In particular, we first review a pragmatic construction based on bit-interleaved coded modulation (BICM) [2], which relies on the underlying binary code to achieve diversity. Secondly, we consider the concatenation of a coded modulation scheme with an inner code that is linear in the field of complex numbers (linear dispersion (LD) code [3]). In both cases, STCs showing full-diversity performance of any desired spectral efficiency are constructed by suitably expanding the signal constellation. In the first case, we have constellation expansion in the complex plane (Ungerboeck's style expansion), while in the second, we have multidimensional expansion induced by the inner code. The aim of this paper is to provide a comparison of these two approaches under low-complexity message passing decoding. We show that multidimensional expansion yields a small performance advantage over complex-plane expansion, while significantly increasing the decoding complexity and peak-to-average power ratio.

II. PRAGMATIC SPACE-TIME CODES

We consider natural STCs (NSTC) coupled with BICM as a pragmatic way to construct *good* STCs (see e.g. [4]). We

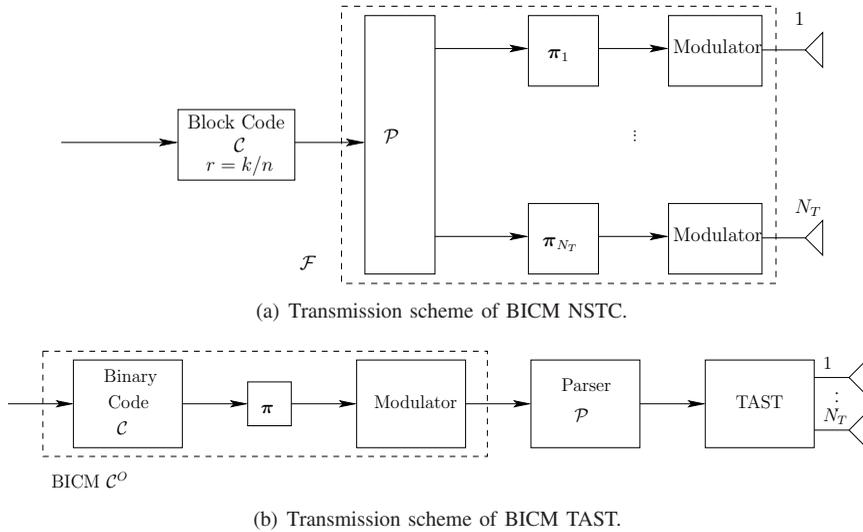


Fig. 1. Block diagrams of the two families of space-time codes.

nickname such scheme BICM NSTC. Such codes are formally defined by a binary block code $\mathcal{C} \subseteq \mathbb{F}_2^N$ of length N and rate r and a spatial modulation function $\mathcal{F} : \mathcal{C} \rightarrow \mathcal{S} \subseteq \mathcal{X}^{N_T \times L}$, such that $\mathcal{F}(\mathbf{c}) = \mathbf{X}$, where $\mathcal{X} \subseteq \mathbb{C}$ is the complex signal constellation. We study the case where \mathcal{F} is obtained as the concatenation of a block/antenna parsing function $\mathcal{P} : \mathbb{Z}_+ \rightarrow \mathbb{Z}_+^2$ such that $\mathcal{P}(n) = (t, \ell)$, $1 \leq n \leq N$, $1 \leq t \leq N_T$, $1 \leq \ell \leq LM$ partitions a codeword $\mathbf{c} \in \mathcal{C}$ into N_T sub-blocks, and blockwise BICM, where each sub-block is independently bit-interleaved according to π_t , $t = 1, \dots, N_T$, and mapped over the signal set \mathcal{X} according to a labeling rule $\mu : \mathbb{F}_2^M \rightarrow \mathcal{X}$, such that $\mu(b_1, \dots, b_M) = x$, where $M = \log_2 |\mathcal{X}|$ (see Figure 1a) ¹. In this case, $N = N_T LM$. The transmission rate (spectral efficiency) of the resulting STC is $R = r N_T M$ bit/s/Hz.

BICM NSTCs are designed assuming a genie aided decoder that produces observables of the transmitted symbols of one antenna, when the symbols from all other antennas are perfectly removed² [8]. In this way, the channel is decomposed into a set of N_T single-input single-output non-interfering parallel channels (see [5] and references therein). We define the *block diversity* of a STC \mathcal{S} as the blockwise Hamming distance,

$$\delta_\beta \triangleq \min_{\mathbf{X}, \mathbf{X}' \in \mathcal{S}} |\{t \in [1, \dots, N_T] : \mathbf{x}_t - \mathbf{x}'_t \neq 0\}|$$

i.e., the minimum number of nonzero rows of $\mathbf{X} - \mathbf{X}'$. Then, with a genie aided decoder BICM NSTCs can achieve diversity $\delta_\beta N_R$ [5]. Notice that applying BICM within a block, preserves the block diversity of the underlying binary code, since the binary labeling rule μ is a bijective correspondence. Thus, the block diversity of \mathcal{S} is equal to the block diversity of \mathcal{C} after the blockwise parsing operated by \mathcal{P} . For a binary code \mathcal{C} of rate r over \mathbb{F}_2 mapped over N_T independent blocks,

a fundamental upper bound on δ_β is provided by the Singleton bound (SB),

$$\delta_\beta \leq 1 + \lfloor N_T (1 - r) \rfloor \quad (3)$$

Consequently, we will search for codes maximizing δ_β , i.e., achieving the SB for all values of N_T . Equation (3) implies that in order to have full block diversity $r \leq 1/N_T$. Hence, in order to transmit a given desired rate R with $\delta_\beta = N_T$, we must have $M = R$. This fixes the size of the signal constellation \mathcal{X} in the BICM scheme.

On the other hand, the evaluation of the rank diversity of BICM NSTCs can be a very involved task especially for constellations with $M > 1$. In the binary case, it is however possible to verify through the stacking construction theorem [6] whether a BICM NSTC is full-rank. The BICM NSTC has full rank-diversity if and only if for all $a_1, \dots, a_{N_T} \in \mathbb{F}_2$ not all zero, the $K \times N$ matrix

$$\Theta = \bigoplus_{t=1}^{N_T} a_t \mathbf{G}_t \mathbf{\Pi}_t$$

has rank K (\bigoplus indicates addition in the binary field \mathbb{F}_2), where $\mathbf{G}_1, \dots, \mathbf{G}_{N_T} \in \mathbb{F}_2^{K \times N}$ are the binary generator matrices of the binary code \mathcal{C} of rate $K/N_T N$ and $\mathbf{\Pi}_1, \dots, \mathbf{\Pi}_{N_T} \in \mathbb{Z}_+^{N \times N}$ are the N_T permutation matrices obtained by applying the permutations π_1, \dots, π_{N_T} to the columns of the identity matrix \mathbf{I} [6]. Notice also that by definition we have that $d_r \leq \delta_\beta \leq N_T$.

III. LD CONCATENATED CODES

In this section we consider the case where the codewords \mathbf{X} of the STC \mathcal{S} are obtained from the concatenation of an outer coded modulation scheme $\mathcal{C}^O \subseteq \mathcal{X}^n$ of rate r_O and length n with an inner LD code. The inner code is formed by a parser \mathcal{P} , that partitions the codewords $\mathbf{c} \in \mathcal{C}^O$ into sub-blocks $\mathbf{c}[j] = [c_1[j], \dots, c_Q[j]]$, $j = 1, \dots, J$ of length Q , with $J = n/Q$ and by a LD space-time modulation function \mathcal{F} defined by,

$$\mathbf{S}[j] = \mathcal{F}(\mathbf{c}[j]) = \sum_{q=1}^Q (c_q[j] \mathbf{G}_q), \quad (4)$$

¹In the remainder of this letter we shall only consider Gray labeling rules, since they are more efficient in quasistatic channels. This conclusion may be reversed in fully-interleaved channels [4].

²The reader will notice the analogy with the case of decision feedback equalization for frequency selective channels, where correct feedback is assumed to design the equalizer filters.

where $\mathbf{G}_q \in \mathbb{C}^{N_T \times T}$ are the LD code generator matrices. Finally, the overall space-time codeword is given by $\mathbf{X} = [\mathbf{S}[1] \dots \mathbf{S}[J]]$. Equation (1) can be rewritten as a virtual MIMO channel with Q inputs and $N_R^v = N_R T$ outputs as,

$$\mathbf{y}[j] = \mathcal{H}\mathbf{c}[j] + \mathbf{z}[j], \quad j = 1, \dots, J \quad (5)$$

where $\mathcal{H} = [I_T \otimes \mathbf{H}]\mathbf{G} \in \mathbb{C}^{N_R^v \times Q}$ is the equivalent channel matrix, \otimes is the Kronecker product, $\mathbf{G} \in \mathbb{C}^{N_T T \times Q}$ is the suitably reformatted generator matrix of the LD code, $\mathbf{y}[k] = \text{vec}(\mathbf{Y}[k])$, $\mathbf{z}[k] = \text{vec}(\mathbf{Z}[k])$, $N_R^v = N_R T$ is the number of virtual receive antennas, and $\text{vec}(\mathbf{A}) = [\mathbf{a}_1^T \dots \mathbf{a}_L^T]^T$, for a matrix $\mathbf{A} = [\mathbf{a}_1 \dots \mathbf{a}_L]$. We will refer to Q as the number of virtual transmit antennas.

In order to perform a meaningful comparison with BICM NSTC, we choose \mathcal{C}^O to be a standard BICM, i.e., a binary code $\mathcal{C} \in \mathbb{F}_2^M$ of rate r whose bit-interleaved codewords are mapped onto the signal set \mathcal{X} according to the binary labeling rule $\mu: \mathbb{F}_2^M \rightarrow \mathcal{X}$ [2]. In this work we are interested in inner LD codes with full-diversity and full-rate, i.e., transmission of $N_L = \min\{N_T, N_R\}$ symbols per channel use. Orthogonal and quasi-orthogonal STCs codes cannot achieve more than 1 symbol per channel use. Therefore, using such codes as inner codes incurs a significant rate loss for $N_L > 1$. As the inner LD code, we use the recently proposed threaded algebraic space-time (TAST) constellations [7]. We nickname such a transmission scheme as BICM TAST (Figure 1b).

The aforementioned algebraic STCs are based on the threaded layering [8], for which a number of component encoders (or layers) N_L and the component codewords $\mathbf{c}_\ell \in \mathcal{C}_\ell$, $\ell = 1, \dots, N_L$ onto the array $\mathcal{T}_{N_T, N_L, N_L}$ following the (layer, antenna, time) indexing triplet

$$\mathcal{P}(\ell, n) = (\ell, |t + \ell - 1|_{N_T}, t), \quad 1 \leq t \leq T, \quad \ell = 1, \dots, N_L,$$

thus having full spatial and temporal spans. Modulo- k operation is denoted by $|\cdot|_k$. The codewords of \mathcal{C}_ℓ , $\ell = 1, \dots, N_L$ are obtained as the set of vectors $\mathbf{v} = \phi_\ell \mathbf{M}\mathbf{s}$, where \mathbf{M} is a rate one full-diversity linear algebraic rotation, ϕ_ℓ , $\ell = 1, \dots, N_L$ are scalar complex coefficients chosen to be Diophantine numbers that ensure that TAST constellation achieves full diversity with maximum-likelihood (ML) decoding (see [7] for details) and $\mathbf{s} \in \mathcal{X}^{N_T}$. In TAST constellations, $N_L = \min(N_T, N_R)$, $T = N_T$, $Q = N_L N_T$, and ϕ_ℓ , $\ell = 1, \dots, N_L$ are chosen such that $|\phi_\ell| = 1$, $\ell = 1, \dots, N_L$. The transmission rate of the resulting BICM TAST is therefore $R = r N_L M$ bit/s/Hz. Notice that while the components of \mathbf{s} belong to \mathcal{X} , the components of \mathbf{v} , which are sent to the antennas for transmission, belong to a scattered constellation with 2^{MN_T} points since by construction any two different \mathbf{s}, \mathbf{s}' are mapped by the rotation \mathbf{M} into points \mathbf{v}, \mathbf{v}' that differ in all N_T components (see [7] and references therein). Therefore, BICM TAST expands the signal constellation by construction, and of \mathbf{M} is N_T and that ϕ_ℓ , $\ell = 1, \dots, N_L$ are Diophantine numbers [7].

With this *divide and conquer* design, full diversity is always guaranteed by the inner code, while coding gain is left to the outer coded modulation. Furthermore, since full-diversity is ensured by the inner code, the designer has some degrees of freedom in choosing r and M in order to achieve the desired transmission rate, as opposed to the BICM NSTC case.

IV. MESSAGE PASSING DECODING

Because of the pseudo-random bit interleaver present in both schemes, ML decoding of BICM NSTC or BICM TAST has generally unaffordable complexity. We therefore resort to iterative techniques based on a factor graph representation. In analogy to the case of multiuser receivers for CDMA [9], applying the belief propagation (BP) algorithm to the STC dependency graph, yields several receivers that approximate the optimal maximum a posteriori (MAP) detection rule. In particular, exact BP reduces the overall receiver to a MAP soft-input soft-output (SISO) bitwise demodulator and a MAP SISO decoder of \mathcal{C} , that exchange extrinsic information probability messages through the iterations. The log-likelihood ratio (LLR) message at the i -th iteration by the MAP SISO bitwise demodulator for the decoder of \mathcal{C} for BICM NSTC, corresponding to the m -th bit of the constellation symbol transmitted over antenna t at discrete time ℓ , is given by

$$\text{LLR}_{ext}^{(i)}(c_{t,\ell,m} | \mathbf{y}_\ell, \mathbf{H}) = \log \frac{\sum_{\mathbf{x} \in \mathcal{X}_{m=0}^t} p(\mathbf{y}_\ell | \mathbf{x}, \mathbf{H}) \prod_{t'=1}^{N_T} \prod_{\substack{m'=1 \\ m' \neq m, t=t'}}^M P_{ext}^{(i-1)}(c_{t',\ell,m'})}{\sum_{\mathbf{x} \in \mathcal{X}_{m=1}^t} p(\mathbf{y}_\ell | \mathbf{x}, \mathbf{H}) \prod_{t'=1}^{N_T} \prod_{\substack{m'=1 \\ m' \neq m, t=t'}}^M P_{ext}^{(i-1)}(c_{t',\ell,m'})}$$

for $1 \leq m \leq M, 1 \leq t \leq N_T, 1 \leq \ell \leq L$, where $\mathcal{X}_{m=a}^t$ is the set of N_T -dimensional symbols for which the m -th bit of the symbol transmitted over antenna t equal to a , $P_{ext}^{(i)}(c)$ denotes extrinsic (EXT) probability (provided by the SISO decoder of \mathcal{C}) of the coded binary symbol c at the i -th iteration with $P_{ext}^{(0)}(c) = 0.5$, and $p(\mathbf{y}_\ell | \mathbf{x}, \mathbf{H}) \propto \exp(-|\mathbf{y}_\ell - \sqrt{\rho}\mathbf{H}\mathbf{x}|^2)$.

The case of BICM TAST is completely analogous, as it suffices to replace N_T by Q , L by J and \mathbf{H} by \mathcal{H} . Exact BP is of exponential complexity in MN_T for BICM NSTC and $MN_L N_T$ for BICM TAST and it is usually approximated by soft-output sphere decoding techniques (see e.g. [10] for recent results on the subject). The complexity of such decoders critically depends on the dimension of the channel matrix. Since the dimension of $\mathcal{H} \in \mathbb{C}^{N_R N_T \times N_L N_T}$ is larger than the dimension of the original channel matrix \mathbf{H} , in general the decoding complexity of BICM TAST codes will be substantially higher than the decoding complexity of BICM NSTCs.

In this work we also consider lower-complexity receivers based on iterative interference cancellation (IC) and linear filtering. For BICM NSTC, the LLR message to the decoder of \mathcal{C} is given by,

$$\text{LLR}_{ext}^{(i)}(c_{t,\ell,m} | z_{t,\ell}^{(i)}, \mathbf{H}) = \log \frac{\sum_{x \in \mathcal{X}_{m=0}} p(z_{t,\ell}^{(i)} | x, \mathbf{H}) \prod_{\substack{m'=1 \\ m' \neq m}}^M P_{ext}^{(i-1)}(c_{t,\ell,m'})}{\sum_{x \in \mathcal{X}_{m=1}} p(z_{t,\ell}^{(i)} | x, \mathbf{H}) \prod_{\substack{m'=1 \\ m' \neq m}}^M P_{ext}^{(i-1)}(c_{t,\ell,m'})}$$

for $1 \leq m \leq M, 1 \leq t \leq N_T, 1 \leq k \leq L$, where now $\mathcal{X}_{m=a}$ is the set of all constellation points of \mathcal{X} with the m -th bit

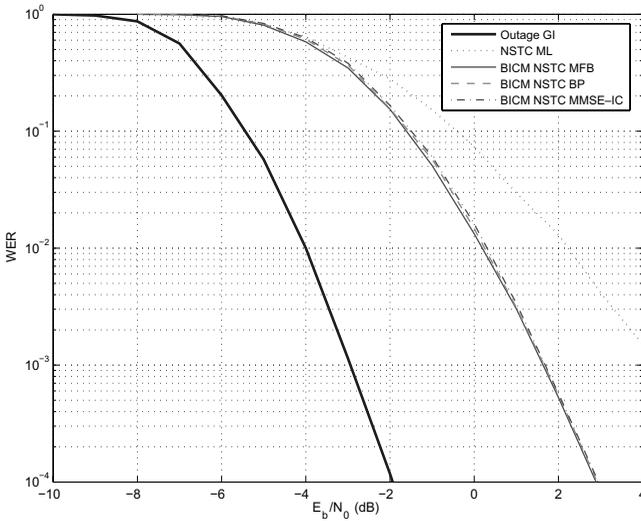


Fig. 2. WER as a function of E_b/N_0 in a MIMO channel with $N_T = 4$, $N_R = 4$ and $R = 1$ bit/s/Hz, with the $(5, 7, 7, 7)_8$ convolutional code, 128 information bits per frame, BPSK modulation and several iterative receivers with 5 decoding iterations.

of the label equal to a , $z_{t,\ell}^{(i)}$ is the output at symbol time ℓ and i -th iteration of the front-end linear filter $\mathbf{f}_t^{(i)}$ of antenna t after IC,

$$z_{t,\ell}^{(i)} = \mathbf{f}_t^{(i)H} \left(\mathbf{y}_\ell - \sqrt{\rho} \sum_{t' \neq t}^{N_T} \mathbf{h}_{t'} \hat{x}_{t',\ell}^{(i-1)} \right),$$

where (dropping antenna and time indexes for simplicity),

$$\hat{x}^{(i)} = \mathbb{E}[x | \text{EXT}] = \sum_{x \in \mathcal{X}} x \prod_{m=1}^M P_{\text{ext}}^{(i)}(c_m)$$

is the minimum mean-square error estimate (conditional mean) of the symbol x given the extrinsic information (briefly denoted by EXT) relative to the bits in the label of x .

In particular, we consider minimum mean squared error (MMSE) IC, for which the filter at the i -th iteration corresponding to the t -th antenna is given by,

$$\mathbf{f}_t^{(i)} = \alpha_t \sqrt{\rho} \mathbf{R}^{-1} \mathbf{h}_t,$$

where $\alpha_t = (\rho \mathbf{h}_t^H \mathbf{R}^{-1} \mathbf{h}_t)^{-1}$ is the normalization constant, $\mathbf{R} = \mathbf{I} + \sqrt{\rho} \sum_{t=1}^{N_T} \mathbf{h}_t \mathbf{h}_t^H v_t$ is the covariance matrix of the input signal to the filter, and $v_t = \mathbb{E}[|x_t - \hat{x}_t|^2]$ is the variance of the residual interference at virtual antenna t (see [9] and references therein). A practical implementation (and in our simulations) we estimate v_t as $v_t \approx 1 - \frac{1}{L} \sum_{\ell=1}^L |\hat{x}_t[\ell]|^2$. Notice that $\mathbf{f}_t^{(i)}$ has to be computed once per virtual transmit antenna and iteration. The proposed algorithm differs from that proposed in [8] in that the latter has to be computed once per symbol interval, transmit antenna and iteration. Obviously, the MMSE IC scheme is also applicable to BICM TAST by using \mathcal{H} instead of \mathbf{H} .

V. EXAMPLES

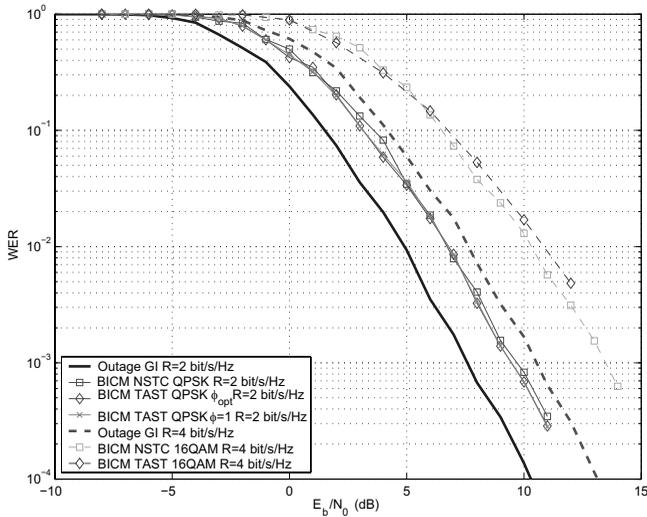
In this section we provide several numerical examples obtained by computer simulation that illustrate the effect of the constellation expansion on the achievable diversity of BICM

NSTC and BICM TAST. For the sake of comparison, we include the outage probability curves with Gaussian inputs (denoted by “GI” in the figures) at the corresponding spectral efficiency.

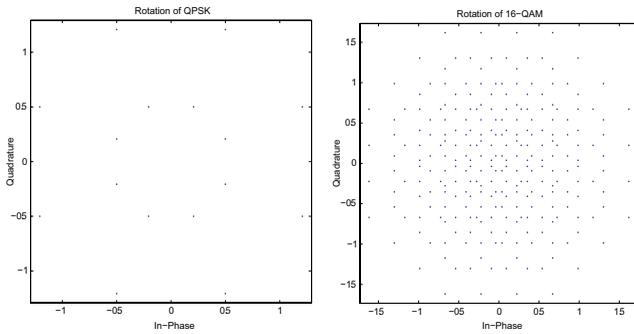
In the case of Figure 2, clearly, the block diversity of \mathcal{C} is $\delta_\beta = 4$. In dotted line we show the word-error rate (WER) for the NSTC with ML decoding. Recall that the NSTC array is constructed using identity permutations [6], and therefore ML decoding is possible using the Viterbi algorithm. The stacking construction theorem yields that the NSTC code is rank deficient. We have also applied the theorem to BICM NSTC with a large number of randomly generated interleaver permutations, and none of them gave a full-rank code. However, as the curves in the figure show, in the WER region of interest (i.e., $10^{-3} - 10^{-4}$) the performance with two suboptimal iterative receivers (BP and MMSE-IC) follows the matched filter bound (MFB), i.e., the performance of the genie aided receiver that has full diversity by definition. Eventually at large SNR and much lower WER, the rank deficiency of the code will show its effect, and the slope of the curves with iterative decoding will change, in a way similar to the different behavior of the waterfall and floor regions in the error curve of turbo-codes. This simple example shows the key role that the block diversity plays in BICM NSTC. We shall say that a scheme shows *full diversity behavior* if the slope of the error curve in a region of interest coincide with the optimal (full-diversity) slope, given by the information outage probability with Gaussian inputs.

In Figure 3 the block diversity of \mathcal{C} is $\delta_\beta = N_T = 2$. In fact, BICM NSTC shows full diversity behavior. On the other hand, BICM TAST achieves full diversity and the coded modulation yields only some coding gain (horizontal shift of the error curve). Figure 3 shows also the scatter diagrams of the TAST constellations. Notice that, in such concatenated scheme, we may set $\phi_\ell = 1$, $\ell = 1, \dots, N_L$ without any noticeable difference in performance, since the outer code removes most of the rank-deficient error events of the inner code. In this way, it suffices to find a good rotation matrix \mathbf{M} in order to construct efficient BICM TAST codes. However, in this case the BICM TAST scheme does not probably achieve full diversity, and shows just the full diversity behavior in the range of WER we have simulated.

Figure 4 clearly illustrates the effect of constellation expansion to achieve full diversity. In order to achieve $R = 3$ bit/s/Hz with QPSK, we need the rate of \mathcal{C} be $r = 3/4$. As we observe, under such configuration BICM TAST achieves full diversity due to its inherent multidimensional constellation expansion. If $\mathbf{M} = \mathbf{I}$ no multidimensional expansion occurs (the transmitted signal constellation at each antenna is QPSK), and full-diversity is obviously not achieved. On the other hand, the diversity of BICM NSTC is governed by the Singleton bound (which in this case yields $\delta_\beta = 1$) and therefore under this configuration it does not achieve full-diversity. However, $R = 3$ bit/s/Hz can also be achieved by using a rate $r = 1/2$ code (which has $\delta_\beta = N_T = 2$) and expanding the signal constellation in the complex plane, i.e., using 8-PSK instead of QPSK. In this case, BICM NSTC achieves a full diversity behavior but it pays about 0.8 dB penalty in average SNR for the expansion with respect to the BICM TAST. The curves



(a)



(b)

(c)

Fig. 3. (a) WER performance in a $N_T = 2$ and $N_R = 2$ MIMO channel with the 4 states $(5, 7)_8$ convolutional code of $r = 1/2$ with 128 information bits per frame, QPSK, 16-QAM modulations with Gray mapping and 5 iterations of BP decoding. The spectral efficiencies are $R = 2$ bit/s/Hz (solid lines) and $R = 4$ bit/s/Hz (dashed lines) respectively. Scatter diagrams of the transmitted TAST constellations for QPSK and (b) 16-QAM with (c) the optimal 2×2 complex rotation [7, Eq. (15)].

for $R = 4$ bit/s/Hz have a similar behavior (for the sake of clarity we omit the curves for BICM NSTC with $r = 2/3$ 8-PSK and BICM TAST with $M = I$ for which full diversity is not achieved). In this case, BICM NSTC and BICM TAST are very close.

Figure 5 reports again the $N_T = 4$ and $N_R = 4$ case with higher spectral efficiencies. We also plot the simulated matched filter bound. In this example we observe that a new effect arises, namely, for too large spectral efficiency, even if the transmission schemes ensure full diversity, the MMSE-IC decoder is not able to remove the interference and achieve the correct slope. The characterization of the thresholds of the spectral efficiency for which the MMSE-IC is able to perform close to ML or BP appears to be a difficult problem and at present there is no satisfactory explanation. In [11] we derived a semi-analytical method based on density evolution and bounding techniques, which however is as complex as simulation due to the outer expectation over the quasi-static fading.

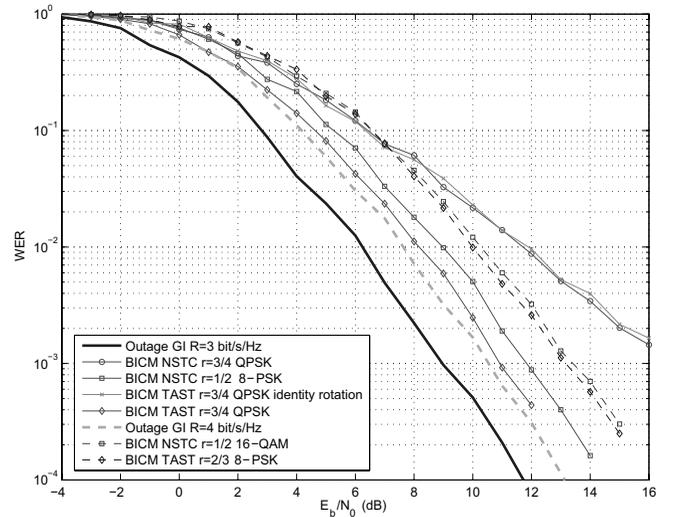


Fig. 4. WER performance in a $N_T = 2$ and $N_R = 2$ MIMO channel with the 4 states convolutional codes and QPSK, 8-PSK and 16-QAM modulations with Gray mapping and 5 iterations of BP decoding for overall spectral efficiencies of $R = 3$ bit/s/Hz (solid lines, with 132 information bits per frame) and $R = 4$ bit/s/Hz (dashed lines, with 128 information bits per frame).

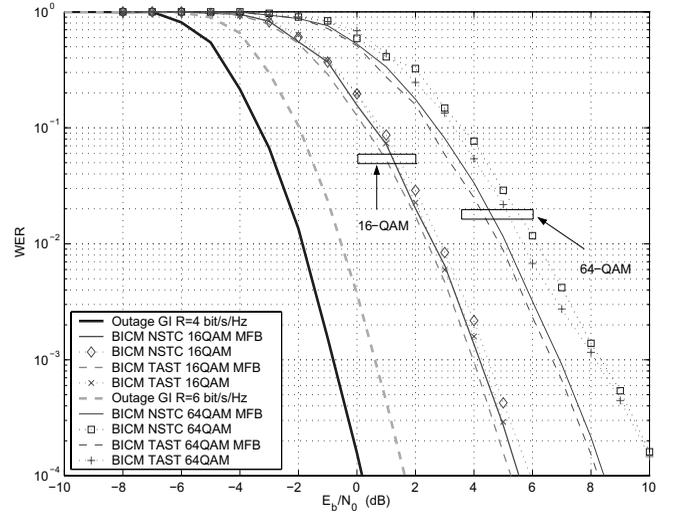


Fig. 5. WER performance of BICM NSTC and BICM TAST in a MIMO channel with $N_T = 4$ and $N_R = 4$, with the 4 states $(5, 7, 7, 7)_8$ convolutional code of $r = 1/4$, using 16 and 64-QAM modulations with Gray mapping and 5 iterations of MMSE-IC decoding. The corresponding spectral efficiencies are $R = 4$ bit/s/Hz (128 information bits per frame) and $R = 6$ bit/s/Hz (120 information bits per frame).

VI. CONCLUSIONS

In this letter we have illustrated the effect of signal constellation expansion on the achievable diversity in quasistatic MIMO channels. In particular we have compared complex plane expansion and expansion due to multidimensional full-diversity rotations. We have shown that under message passing decoding, concatenated LD STCs enjoy higher design flexibility in the choice of the outer binary coding rate and of the underlying constellation size, but yield only small performance advantage with respect to the simpler BICM NSTCs, while considerably increasing the decoding complexity and the peak-to-average power ratio. Therefore, in practical applications such as IEEE802.11n, considering BICM NSTC (with very powerful binary outer codes) is fully justified.

REFERENCES

- [1] V. Tarokh, N. Seshadri, and A. R. Calderbank, "Space-time codes for high data rate wireless communication: performance criterion and code construction," *IEEE Trans. Inf. Theory*, vol. 44, no. 2, pp. 744–765, March 1998.
- [2] G. Caire, G. Taricco, and E. Biglieri, "Bit-interleaved coded modulation," *IEEE Trans. Inf. Theory*, vol. 44, no. 3, pp. 927–946, May 1998.
- [3] B. Hassibi and B. Hochwald, "High-rate codes that are linear in space and time," *IEEE Trans. Inf. Theory*, vol. 48, no. 7, pp. 1804–1824, July 2002.
- [4] J. Boutros, N. Gresset, and L. Brunel, "Turbo coding and decoding for multiple antenna channels," in *Proc. 3th Int. Symp. Turbo Codes and Applications 2003*.
- [5] R. Knopp and P. Humblet, "On coding for block fading channels," *IEEE Trans. Inf. Theory*, vol. 46, no. 1, pp. 1643–1646, July 1999.
- [6] R. Hammons and H. El Gamal, "On the theory of space-time codes for PSK modulation," *IEEE Trans. Inf. Theory*, vol. 46, no. 2, pp. 524–542, Mar. 2000.
- [7] M. O. Damen and H. El Gamal, "Universal space-time codes," *IEEE Trans. Inf. Theory*, vol. 49, no. 5, pp. 1059–1063, May 2003.
- [8] H. El Gamal and R. Hammons, "A new approach to layered space-time coding and signal processing," *IEEE Trans. Inf. Theory*, vol. 47, no. 6, pp. 2321–2334, Sept. 2001.
- [9] J. Boutros and G. Caire, "Iterative joint decoding: unified framework and asymptotic analysis," *IEEE Trans. Inf. Theory*, vol. 48, no. 7, pp. 1772–1793, July 2002.
- [10] J. Boutros, N. Gresset, L. Brunel, and M. Fossorier, "Soft-input soft-output lattice sphere decoder for linear channels," in *Proc. IEEE Global Communications Conference 2003*.
- [11] A. Guillén i Fàbregas and G. Caire, "Analysis and design of natural and threaded space-time codes with iterative decoding," in *Proc. 36th Asilomar Conference 2002*.