A Class of Low Complexity Iterative Equalizers for Space-Time BICM over MIMO Block Fading Multipath AWGN Channel

Sami Chtourou, Student Member, IEEE, Raphaël Visoz, Member, IEEE, and Antoine O. Berthet, Member, IEEE

Abstract— **In this paper, we propose a class of iterative equalizers for space-time bit interleaved coded modulation over MIMO block fading multipath AWGN channel. Simulations show that the proposed receivers are able to reach the matched-filter bound and that the tested space-time coding schemes have the potential to get close to the Shannon theoretical limit.**

Keywords— **Space-Time BICM, MIMO Block Fading Multipath AWGN Channel, Interference Cancellation, Iterative Decoding**

I. Introduction

It was first conjectured in [1] that Space-Time Bit-Interleaved Coded Modulation (STBICM) is a simple and appealing signaling scheme for ergodic Multiple-Input Multiple-Output (MIMO) flat fading channel. Under iterative decoding (ID) assumption, this scheme has the potential to perform very close to the Shannon theoretical limit [2]. We would expect this good behavior to encompass the non-ergodic case as well. Several recent contributions have all confirmed, at least by Monte-Carlo simulations, the attractiveness of STBICM-ID for MIMO block fading AWGN channel with InterSymbol Interference (ISI) [3]. However, the main drawback of STBICM-ID lies in the complexity involved at the receiver side. As well known, optimal joint Maximum A Posteriori (MAP) criterion is purely infeasible but can be approached by separating the MIMO ISI detection and the outer decoding (both performed in symbol-by-symbol MAP sense) and by exchanging randomized soft information between them in an iterative fashion [4] [5]. Unfortunately, the complexity of MAP ISI detection, being directly related to the number of states in the channel trellis, increases exponentially with transmission rate and channel selectivity and explodes in the MIMO case. Some previous works have attempted to tackle this complexity impediment by resorting to massive trellis state reduction together with Per Survivor Processing (PSP) ([6] and the references therein). However, for large MIMO systems and/or high-order modulations, the number of transitions per state becomes the true limiting factor of such techniques. On the other hand, soft interference cancellation algorithms based on the recursive calculation of a MMSE filter taking into account prior statistics on decisions have received a great amount of attention since the last six years, revealing a powerful alternative to MAP ISI detection (and its suboptimal versions). Various methods have been proposed to equalize Single-Input Single-Output (SISO) channels with memory or to detect multiple users in CDMA systems [7] [8] [9] [10]. Exploiting the analogy between time domain and spatial domain spreading, some recent contributions have revisited the Wang et al.'s equations in the context of multiuser MIMO channel with memory [11] [12].

This paper describes a new class of efficient equalizers for STBICM transmitted over MIMO block-fading multipath channel. Our objective is to solve the problem of Inter-Symbol Interference (ISI) and Multi-Antenna Interference (MAI) with minimal complexity, i.e., polynomial in all system parameters, while performing as close as possible from the Shannon theoretical limit. In Section II, we introduce the communication model. In section III, we briefly derive the well-known Wang et al.'s approach in the context of MIMO transmission, where ISI equalization and MAI resolution are jointly performed for all individual antennas. In section IV, we propose an alternative approach, in which ISI and MAI cancellation tasks are decoupled, leading to a much lower computational complexity. Section V is devoted to Monte-Carlo simulation results and comments. We conclude the paper in section VI, opening future research topics.

Notation

- The superscripts *,$^\mathsf{T}$ and † indicate conjugate, transpose and Hermitian transpose, respectively.
- diag{.}, tr{.}, det{.} denote diagonal, trace, and determinant operators on square matrices, respectively.
- Probability density functions (pdf) are denoted $p\,(.)$ and probability mass functions (pmf) are denoted $\Pr\,[.]$.

II. Communication model

We consider a MIMO B-block fading multipath AWGN channel with T transmit and R receive antennas and memory M [13] [14]. Channel State Information (CSI) is perfectly known at the receiver and unknown at the transmitter. Fading blocks are thought as separated both in time and frequency and may be correlated or not. This model is well suited to represent a slowly time-varying MIMO multipath channel where blocks may either result from frequency-hopping in TDMA systems or be identified with subcarriers in OFDM systems. In the simulation Sec-

Sami Chtourou and Raphaël Visoz are currently with France Telecom R&D, 38-40 rue du Général Leclerc, 92794 Issy Les Moulineaux, France. Antoine O. Berthet is currently with Ecole Supérieure d'Electricité (SUPELEC), Wireless Communications Department, Plateau de Moulon, 91192 Gif-sur-Yvette, France.

tion, the block-static subcase (i.e., $B = 1$) benefits from particular attention.

A. Space-time bit-interleaved coded modulation

Let $\mathcal{C}$ be a linear code of length N and rate ρ over $\mathbb{F}_2$. Produced code words $\mathbf{c} = \{c_n\} \in \mathbb{F}_2^N$ enter a well-designed bit interleaver $\mathcal{I}$[1], whose matrix output $\mathbf{D} \in \mathbb{F}_2^{B \times Q \times L}$ is segmented into B matrices $\mathbf{D}^b \in \mathbb{F}_2^{Q \times L}$, $b = 1, ..., B$. Columns of matrices $\mathbf{D}^b$ are vectors $\mathbf{d}_k^b \in \mathbb{F}_2^Q$, $k = 1, ..., L$, referred to as "vector symbol digit", containing T sub-vectors $\mathbf{d}_{t,k}^b \in \mathbb{F}_2^{Q_t}$, $t = 1, ..., T$ (one per channel input), with stacked binary components. Clearly, $Q = \sum_t Q_t$. We access to the i^{th} binary component of $\mathbf{d}_{t,k}^b$ via the dedicated notation $d_{\langle t,i \rangle,k}^b$. Within each matrix $\mathbf{D}^b$, all subvectors $\mathbf{d}_{t,k}^b$ are mapped onto the corresponding complex signal set $\mathcal{A}_t$ of cardinality $|\mathcal{A}_t| = 2^{Q_t}$ through a labeling application $\phi_t : \{0,1\}^{Q_t} \rightarrow \mathcal{A}_t$. After signal mapping, this coding-modulation process can be equivalently regarded as a space-time modulation-coding scheme where every code word $\mathbf{X}$ contains $B \times T \times L$ complex symbols divided into B distinct matrices $\mathbf{X}^b$, $B = 1, ..., B$, whose columns $\mathbf{x}_k^b \in \mathcal{A}_1 \times ... \times \mathcal{A}_T \subset \mathbb{C}^T$, $k = 1, ..., L$ are referred to as "vector constellation symbol". The i^{th} digit of subvector $\mathbf{d}_{t,k}^b$ can be recovered from $\mathbf{x}_k^b$ by simple reverse mapping $\phi_{t,i}^{-1}$. It can be convenient to define the global labeling application $\phi : \{0,1\}^Q \rightarrow \mathcal{A}_1 \times ... \times \mathcal{A}_T$ which maps vectors symbol digit $\mathbf{d}_k^b$ into vectors constellation symbol $\mathbf{x}_k^b$. The i^{th} digit of vector $\mathbf{d}_k^b$ can be recovered from $\mathbf{x}_k^b$ by simple reverse mapping ϕ_i^{-1}. Falling into the general class of space-time codes, this architecture offers a spectral efficiency in bits per channel use (p.c.u) :

$$\eta = \rho \sum_{t=1}^{T} Q_t \tag{1}$$

under ideal Nyquist band-limited filtering assumption. Without loss of generality, we assume that sets $\mathcal{A}_t$ are identical for all transmit antennas and that per-constellation Gray labeling is used[2].

B. MIMO block fading multipath channel

Let $\mathbf{H}^b \in \mathbb{C}^{R \times T \times (M+1)}$ denote the MIMO fading multipath channel block $b = 1, ..., B$ and $\mathbf{H} = \{\mathbf{H}^b\}$ the collection of all channel blocks. Let also $\mathbf{X}^b \in \mathbb{C}^{T \times L}$ and $\mathbf{Y}^b \in \mathbb{C}^{R \times L}$ be the "matrix constellation symbol" and the "matrix channel output". The discrete-time base-band equivalent vector channel output $\mathbf{y}_k^b \in \mathbb{C}^R$ at time $k = 1, ..., L$ can be written as

$$\mathbf{y}_k^b = \sum_{m=0}^{M} \mathbf{H}_m^b \mathbf{x}_{k-m}^b + \mathbf{w}_k^b, \tag{2}$$

where $\mathbf{x}_k^b \in \mathbb{C}^T$ are the vector constellation symbol transmitted at time k, $\mathbf{H}_m^b \in \mathbb{C}^{R \times T}$ is the m^{th} matrix tap of the channel impulse response, $\mathbf{w}_k^b \in \mathbb{C}^R$ is the vector of additive complex noise. The vectors of additive complex noise $\mathbf{w}_k^b$ are assumed zero-mean independent identically distributed (i.i.d) circularly symmetric complex Gaussian and thus follow the pdf $\mathcal{N}\left(\mathbf{0}, \sigma^2 \mathbf{I}\right)$. Channel $\mathbf{H}^b$, constant along the corresponding block duration, has a Finite Impulse Response (FIR) of length $M+1$, whose symbol-spaced taps $\mathbf{H}_0^b, ..., \mathbf{H}_M^b$ are $R \times T$ complex random matrices with zero-mean and mean power satisfying the normalization constraints

$$\mathbb{E}\left[\operatorname{diag}\left\{\sum_{m=0}^{M} \mathbf{H}_m^b \mathbf{H}_m^{b\dagger}\right\}\right] = T\mathbf{I} \tag{3}$$

in the case of an equal power system. Uniform power allocation is a fair policy when no CSI is available at the transmitter. For subsequent derivations, it is convenient to introduce the equivalent length-L_F sliding window model ($L_F = L_1 + L_2 + 1$) for the received signal

$$\underline{\mathbf{y}}_k^b = \underline{\mathbf{H}}^b \underline{\mathbf{x}}_k^b + \underline{\mathbf{w}}_k^b \tag{4}$$

where $\underline{\mathbf{x}}_k^b, \underline{\mathbf{y}}_k^b, \underline{\mathbf{w}}_k^b$ are stacked vectors

$$\begin{aligned} \underline{\mathbf{x}}_k^b &= \left[\mathbf{x}_{k+L_1}^{b\mathsf{T}} \ \cdots \ \mathbf{x}_{k-L_2-M}^{b\mathsf{T}}\right]^{\mathsf{T}} \\ \underline{\mathbf{y}}_k^b &= \left[\mathbf{y}_{k+L_1}^{b\mathsf{T}} \ \cdots \ \mathbf{y}_{k-L_2}^{b\mathsf{T}}\right]^{\mathsf{T}} \\ \underline{\mathbf{w}}_k^b &= \left[\mathbf{w}_{k+L_1}^{b\mathsf{T}} \ \cdots \ \mathbf{w}_{k-L_2}^{b\mathsf{T}}\right]^{\mathsf{T}} \end{aligned} \tag{5}$$

and where $\underline{\mathbf{H}}^b$ is the Sylvester channel matrix

$$\underline{\mathbf{H}}^b = \begin{bmatrix} \mathbf{H}_0^b & \cdots & \mathbf{H}_M^b & & \\ & \ddots & \cdots & \ddots & \\ & & \mathbf{H}_0^b & \cdots & \mathbf{H}_M^b \end{bmatrix} \tag{6}$$

of dimension $RL_F \times T\left(L_F + M\right)$.

C. Information-theoretic aspects

When the code word spans over a limited number of channel blocks, only the cumulative distribution (outage probability) of the channel capacity is meaningful [15]. Given a transmission rate η, the outage probability of our model is formally defined as:

$$P_{out}\left(\eta\right) = \Pr\left(\frac{1}{B}\sum_{b=1}^{B} C(\mathbf{H}^b) < \eta\right) \tag{7}$$

where $C(\mathbf{H}^b)$ is the channel capacity associated to $\mathbf{H}^b$ given by:

$$C(\mathbf{H}^b) = \int_0^1 \log_2\left(\det\left\{\mathbf{I} + \frac{\gamma}{T}\mathbf{H}^b(\theta)\mathbf{H}^b(\theta)^\dagger\right\}\right) d\theta \tag{8}$$

[1] Refer to [13] for practical details on the design of good interleavers for block fading channel.

[2] Labeling strategies have received particular attention since the last four years. Although natural labeling is often considered as the best choice for STBICM-ID transmitted over scalar and MIMO flat-fading channels, we found that this conclusion is no longer true for frequency-selective channels and that Gray labeling turns out to be the best compromise between the first and the last iteration performance.

with γ the average SNR per receive antenna and $\mathbf{H}^b(\theta)$ the Discrete Fourier Transform (DFT) associated with the b^{th} MIMO channel FIR, i.e,

$$\mathbf{H}^b(\theta) = \sum_{m=0}^{M} \mathbf{H}_m^b e^{-j2\pi m\theta} \tag{9}$$

As well known, the capacity is achieved when $\mathbf{x}$ is a circularly symmetric complex Gaussian with zero mean and covariance $\frac{\gamma}{T}\mathbf{I}$. If $\mathbf{x}$ belongs to the discrete alphabet $\mathcal{A}_1 \times ... \times \mathcal{A}_T \subset \mathbb{C}^T$, the maximum achievable rate should be normally given by the average mutual information $I(\mathbf{x};\mathbf{y})$. This leads to the definition of an information outage probability

$$P_{out}(\eta) = \Pr\left(\frac{1}{B}\sum_{b=1}^{B} I^b(\mathbf{x};\mathbf{y}) < \eta\right) \tag{10}$$

which, in practice, is quite close to the outage probability for the spectral efficiencies presented in section V[3].

III. Block-iterative joint ISI and MAI cancellation

Proposed equalizers process each received data block $\mathbf{Y}^b$, $b = 1, ..., B$ separately, in an iterative fashion, making use of refined probabilistic information fed back by the outer decoder. All described signals and devices should be block-indexed. We suppress this dependency to simplify the notations.

A. Joint ISI and MAI cancellation

Considering antennas as "distinct users", the direct approach solves the classical problem of multiuser detection in the presence of multiple paths [8, section V]. The equalizer is made of a bank of T one-dimensional MMSE filters of dimension $1 \times RL_F$, each of them being followed by a SISO detector. We now focus on antenna t. At iteration l, the vector of soft MMSE symbol estimates used to softly regenerate the MAI and ISI corrupting $x_{t,k}$ is

$$\underline{\widetilde{\mathbf{x}}}_{t,k}^l = \left[\widetilde{x}_{1,k+L_1}^l, ..., \widetilde{x}_{t-1,k}^l, 0, \widetilde{x}_{t+1,k}^l, ..., \widetilde{x}_{T,k-L_2-M}^l\right]^\mathsf{T} \tag{11}$$

Given logarithmic priors $\mathcal{P}_{\langle t,i\rangle,k}^l$ on all symbol digits, the soft MMSE symbol estimates $\widetilde{x}_{t,k}^l$ are obtained as

$$\widetilde{x}_{t,k}^l = \sum_{x\in\mathcal{A}_t} x \frac{\exp\left\{\sum_{i=1}^{Q_t}\phi_{t,i}^{-1}(x)\,\mathcal{P}_{\langle t,i\rangle,k}^l\right\}}{\prod_{i=1}^{Q_t}\left[1+\exp\left\{\mathcal{P}_{\langle t,i\rangle,k}^l\right\}\right]} \tag{12}$$

At first iteration, logarithmic priors are equal to 0. After soft interference cancellation, the input vector at time k for the MMSE filter is expressed as

$$\underline{\widetilde{\mathbf{y}}}_{t,k}^l = \underline{\mathbf{y}}_k - \underline{\mathbf{H}}\underline{\widetilde{\mathbf{x}}}_{t,k}^l \tag{13}$$

[3]Note that the definition of outage or information outage probabilities assumes infinitely long code words. Since, in practice, the BLER performance of our codes actually depends on the chosen length N, the comparison might be somehow unfair. This problem is currently under investigation and exceeds the scope of the paper.

The (biased) filter $\mathbf{f}_t^l$ which minimizes the unconditional MSE

$$\left\|x_{t,k} - \mathbf{f}_t^l\underline{\widetilde{\mathbf{y}}}_{t,k}^l\right\|^2 \tag{14}$$

is easily obtained from the projection theorem as

$$\mathbf{f}_t^l = \mathbf{e}_{\Delta(t)}^\dagger\underline{\mathbf{H}}^\dagger\left[\underline{\mathbf{H}}\mathbf{\Theta}_t^l\underline{\mathbf{H}}^\dagger + \sigma^2\mathbf{I}\right]^{-1} \tag{15}$$

considering the stochastic innerproduct $\langle\mathbf{x},\mathbf{y}\rangle = \mathbb{E}\left\{\mathbf{x}\mathbf{y}^\dagger\right\}$ with associated norm $\|.\|$. In (15), $\mathbf{e}_{\Delta(t)}$ is a $T(L_F+M)\times 1$ vector with a 1 at position $\Delta(t) = TL_1 + t$ and 0's elsewhere. Assuming sufficiently large space-time interleaving, the covariance matrix $\mathbf{\Theta}_t^l$ is expressed as

$$\mathbf{\Theta}_t^l \triangleq \mathbb{E}\left\{\left(\underline{\mathbf{x}}_k - \underline{\widetilde{\mathbf{x}}}_{t,k}^l\right)\left(\underline{\mathbf{x}}_k - \underline{\widetilde{\mathbf{x}}}_{t,k}^l\right)^\dagger\right\} = \tag{16}$$

$$\mathrm{diag}\left\{1-\mathbb{E}\left\{\left|\widetilde{x}_{1,k+L_1}^l\right|^2\right\}, ..., 1, ..., 1-\mathbb{E}\left\{\left|\widetilde{x}_{T,k-L_2-M}^l\right|^2\right\}\right\} \tag{17}$$

with the 1 located at position $\Delta(t)$. The following equality holds for each term on the diagonal

$$\mathbb{E}\left\{\left|\widetilde{x}_{t',k'}^l\right|^2\right\} \approx \frac{1}{LT}\sum_{k=1}^{L}\widetilde{x}_k^{l\dagger}\widetilde{x}_k^l,\ \forall(t',k') \neq (t,k) \tag{18}$$

At the output of the MMSE filter, we turn to the unbiased model

$$z_{t,k}^l = \frac{\mathbf{f}_t^l}{\mathbf{f}_t^l\underline{\mathbf{H}}\mathbf{e}_{\Delta(t)}}\underline{\widetilde{\mathbf{y}}}_{t,k}^l = x_{t,k} + \zeta_{t,k}^l \tag{19}$$

in which $\zeta_{t,k}^l$ represents the residual interference plus noise. The analytical expression of the unconditional SINR β_t^l is given by

$$\beta_t^l = \frac{\mathbf{f}_t^l\underline{\mathbf{H}}\mathbf{e}_{\Delta(t)}}{1-\mathbf{f}_t^l\underline{\mathbf{H}}\mathbf{e}_{\Delta(t)}} \tag{20}$$

Processing $z_{t,k}^l$ from (19) as the output of a virtual SISO AWGN channel with variance $1/\beta_t^l$, the detector computes logarithmic extrinsic probability ratios on symbol digits $i = 1, ...Q$ as

$$\mathcal{L}_{\langle t,i\rangle,k}^{l,\det} = \ln\frac{\sum_{x\in\mathbb{X}_{t,i}^1} p\left(z_{t,k}^l\,|x\right)\exp\left\{\sum_{j\in\mathbb{J}_{t,i}}\phi_{t,j}^{-1}(x)\mathcal{P}_{\langle t,j\rangle,k}^l\right\}}{\sum_{x\in\mathbb{X}_{t,i}^0} p\left(z_{t,k}^l\,|x\right)\exp\left\{\sum_{j\in\mathbb{J}_{t,i}}\phi_{t,j}^{-1}(x)\mathcal{P}_{\langle t,j\rangle,k}^l\right\}} \tag{21}$$

with

$$\begin{array}{c}\mathbb{X}_{t,i}^\varepsilon = \left\{x\in\mathcal{A}_t\,\left|\phi_{t,i}^{-1}(x) = \varepsilon\right.\right\}\\ \mathbb{J}_{t,i} = \left\{j\in 1, ..., Q_t\,|j\neq i\right\}\end{array} \tag{22}$$

B. Iterative scheduling

Logarithmic extrinsic probability ratios $\mathcal{L}_{\langle t,i\rangle,k}^{l,\det}$ are collected over all blocks $b = 1, ..., B$ and rearranged, after space-time deinterleaving $\mathcal{I}^{-1}$, into one single vector of log ratios $\mathcal{Y}_n^l$, $n = 1, ..., N$, on the basis of which logarithmic extrinsic probability ratios on code word bits can be computed as

$$\mathcal{L}_n^{l,\text{dec}} = \ln \frac{\sum_{\mathbf{c}\in\mathcal{C}|c_n=1} \exp\left\{\sum_{n'\neq n} c_{n'}\mathcal{Y}_{n'}^l\right\}}{\sum_{\mathbf{c}\in\mathcal{C}|c_n=0} \exp\left\{\sum_{n'\neq n} c_{n'}\mathcal{Y}_{n'}^l\right\}} \tag{23}$$

After space-time reinterleaving $\mathcal{I}$, the vector of $\mathcal{L}_n^l$ is scattered into B real-valued matrices of dimension $QT \times L$. Logarithmic extrinsic probability ratios $\mathcal{L}_n^{l,\text{dec}}$ from the outer decoder at iteration l serve as logarithmic priors $\mathcal{P}_{\langle t,i\rangle,k}^{l+1}$ on symbol digits for the equalizer at iteration $l+1$. This first approach is naturally suited to the multiuser MIMO channel model. It requires T matrix inversions in $O(R^3 L_F^3)$ (one for each antenna).

IV. Decoupling MIMO ISI and MAI cancellation

We now propose to see vector constellation symbol $\mathbf{x}_k$ as belonging to a T-dimensional modulation and to derive one single multidimensional MMSE filter $\mathbf{F}^l$ to cancel the ISI created by other T-dimensional symbols $\mathbf{x}_i$, $i \neq k$. Once the ISI has been (partially) removed, we detect symbol digits in $\mathbf{x}_k$ just as if the MIMO channel was flat. This becomes a quite appealing MIMO detection scenario for which efficient algorithms have already been proposed in the literature. Not only this approach offers obvious computational savings, but it also differs conceptually. If ISI cancellation still relies on MMSE criterion, a new freedom degree appears for the one employed in MIMO detection. Different types of algorithms can be used ranging from the MAP MIMO detector to the very suboptimal SUMF-IC MIMO detector. Prominent among them is the list-APP sphere-decoder which implements the MAP criterion with polynomial complexity in system parameters. Because of the closed connections between all devices, this freedom degree impacts on the behavior of the global iterative algorithm (convergence speed and asymptotic performance) and allows to adapt the receiver complexity to the system load or to the channel selectivity.

A. Multidimensional ISI cancellation

The equalizer is made of a single multidimensional MMSE filter of dimension $T \times RL_F$ followed by a MIMO detector. At iteration l, the vectors of soft MMSE vestor symbol estimates used to softly regenerate the multidimensional ISI corrupting $\mathbf{x}_k$ is

$$\underline{\widetilde{\mathbf{x}}}_k^l = \left[\widetilde{\mathbf{x}}_{k+L_1}^{l\mathsf{T}}, ..., \widetilde{\mathbf{x}}_{k+1}^{l\mathsf{T}}, \mathbf{0}^\mathsf{T}, \widetilde{\mathbf{x}}_{k-1}^{l\mathsf{T}}, ..., \widetilde{\mathbf{x}}_{k-L_2-M}^{l\mathsf{T}}\right]^\mathsf{T} \tag{24}$$

Given logarithmic priors $\mathcal{P}_{i,k}^l$ on all symbol digits, the vector symbol estimates $\widetilde{\mathbf{x}}_k^l$ are obtained as

$$\widetilde{\mathbf{x}}_k^l = \sum_{\mathbf{x}\in\mathcal{A}} \mathbf{x} \frac{\exp\left\{\sum_{i=1}^Q \phi_i^{-1}(\mathbf{x})\,\mathcal{P}_{i,k}^l\right\}}{\prod_{i=1}^Q \left[1+\exp\left\{\mathcal{P}_{i,k}^l\right\}\right]} \tag{25}$$

At first iteration, logarithmic priors are equal to 0. After soft interference cancellation, the input vector for the MMSE filter is expressed as

$$\underline{\widetilde{\mathbf{y}}}_k^l = \underline{\mathbf{y}}_k - \underline{\mathbf{H}}\underline{\widetilde{\mathbf{x}}}_k^l \tag{26}$$

The multidimensional (biased) filter $\mathbf{F}^l$ which minimizes the unconditional MSE

$$\left\|\mathbf{x}_k - \mathbf{F}^l \underline{\widetilde{\mathbf{y}}}_k^l\right\|^2 \tag{27}$$

is similarly obtained from the projection theorem as

$$\mathbf{F}^l = \mathbf{E}_\Delta^\dagger \underline{\mathbf{H}}^\dagger \left[\underline{\mathbf{H}}\boldsymbol{\Theta}^l\underline{\mathbf{H}}^\dagger + \sigma^2\mathbf{I}\right]^{-1} \tag{28}$$

again considering the stochastic innerproduct $\langle \mathbf{x}, \mathbf{y}\rangle = \mathbb{E}\left\{\mathbf{x}\mathbf{y}^\dagger\right\}$ with associated norm $\|.\|$. In (28), $\mathbf{E}_\Delta^\dagger$ is the $T \times T(L_F+M)$ matrix

$$\mathbf{E}_\Delta^\dagger = [\underbrace{\cdots\ \mathbf{0}\ \cdots}_{L_1}\mathbf{I}\underbrace{\cdots\ \mathbf{0}\ \cdots}_{L_2+M}] \tag{29}$$

Assuming sufficiently large space-time interleaving, the covariance matrix $\boldsymbol{\Theta}^l$ is expressed as

$$\boldsymbol{\Theta}^l \triangleq \mathbb{E}\left\{\left(\underline{\mathbf{x}}_k - \underline{\widetilde{\mathbf{x}}}_k^l\right)\left(\underline{\mathbf{x}}_k - \underline{\widetilde{\mathbf{x}}}_k^l\right)^\dagger\right\} = \tag{30}$$

$$\text{diag}\left\{\mathbf{I} - \mathbb{E}\left\{\widetilde{\mathbf{x}}_{k+L_1}^l\widetilde{\mathbf{x}}_{k+L_1}^{l\dagger}\right\}, ..., \mathbf{I}, ..., \mathbf{I} - \mathbb{E}\left\{\widetilde{\mathbf{x}}_{k-L_2-M}^l\widetilde{\mathbf{x}}_{k-L_2-M}^{l\dagger}\right\}\right\} \tag{31}$$

with identity matrix $\mathbf{I}$ located at position L_1+1. The following equality holds for each term on the diagonal

$$\mathbb{E}\left\{\widetilde{\mathbf{x}}_{k'}^l\widetilde{\mathbf{x}}_{k'}^{l\dagger}\right\} \approx \left(\frac{1}{LT}\sum_{k=1}^L \widetilde{\mathbf{x}}_k^{l\dagger}\widetilde{\mathbf{x}}_k^l\right)\mathbf{I},\ \forall k' \neq k \tag{32}$$

At the output of the MMSE filter, we obtain the biased model

$$\mathbf{z}_k^l = \mathbf{F}^l\underline{\widetilde{\mathbf{y}}}_k^l = \mathbf{F}^l\underline{\mathbf{H}}\mathbf{E}_\Delta\mathbf{x}_k + \boldsymbol{\zeta}_k^l \tag{33}$$

in which $\boldsymbol{\zeta}_k^l$ represents the residual interference plus noise with covariance matrix

$$\mathbb{E}\left\{\boldsymbol{\zeta}_k^l\boldsymbol{\zeta}_k^{l\dagger}\right\} = \left(\mathbf{I} - \mathbf{F}^l\underline{\mathbf{H}}\mathbf{E}_\Delta\right)\mathbf{E}_\Delta^\dagger\underline{\mathbf{H}}^\dagger\mathbf{F}^{l\dagger} \tag{34}$$

If the time correlation has no real impact for subsequent MIMO detection, the spatial correlation plays a key role. Spatial whitening can be efficiently done by a Cholesky factorization of the noise covariance matrix. Assume that $\mathbb{E}\left\{\boldsymbol{\zeta}_k^l\boldsymbol{\zeta}_k^{l\dagger}\right\} = \mathbf{L}\mathbf{L}^\dagger$. Whitening the noise yields the new model

$$\mathbf{z}_k^l = \mathbf{G}^l\mathbf{x}_k + \boldsymbol{\zeta}_k^l \tag{35}$$

in which $\mathbf{G}^l = \mathbf{L}^{-1}\mathbf{F}^l\underline{\mathbf{H}}\mathbf{E}_\Delta$ and $\boldsymbol{\zeta}_k^l$ is now a zero-mean random vector with covariance matrix $\mathbb{E}\left\{\boldsymbol{\zeta}_k^l\boldsymbol{\zeta}_k^{l\dagger}\right\} = \sigma^2\mathbf{I}$. The analytical expression of the unconditional SINR β^l becomes

$$\beta^l = \frac{1}{\sigma^2}\text{tr}\left\{\mathbf{G}^{l\dagger}\mathbf{G}^l\right\} \tag{36}$$

Under ideal decoder feedback assumption, it is easy to prove that the maximum possible global SNR β_{MFB}, given by the Matched Filter Bound (MFB)

$$\beta_{MFB} = \frac{1}{\sigma^2}\mathrm{tr}\left\{\sum_{m=0}^{M}\mathbf{H}_m^\dagger\mathbf{H}_m\right\} \tag{37}$$

is reached. Hence, the proposed turbo-equalizer is asymptotically optimum. This second approach requires only one matrix inversion in $O(R^3L_F^3)$ (instead of T) and a Cholesky factorization in $O(T^3)$ for the spatial whitening of the noise vector at the output of the MIMO ISI canceler. The sphere-decoder requires another Cholesky factorization of the Gram matrix $\mathbf{G}^{l\dagger}\mathbf{G}^l$ (in $O(T^3)$ also).

B. Selected algorithms for MIMO detection

B.1 MAP MIMO detector

Processing $\mathbf{z}_k^l$ from (35) as the output of a virtual MIMO fading flat AWGN channel, the MAP MIMO detector computes logarithmic extrinsic probability ratios on symbol digits $i = 1, ...Q$ as

$$\mathcal{L}_{i,k}^{l,\det} = \ln\frac{\sum_{\mathbf{x}\in\mathbb{X}_i^1} p\left(\mathbf{z}_k^l\left|\mathbf{x}\right.\right)\exp\left\{\sum_{j\in\mathbb{J}_i}\phi_j^{-1}(\mathbf{x})\mathcal{P}_{j,k}^l\right\}}{\sum_{\mathbf{x}\in\mathbb{X}_i^0} p\left(\mathbf{z}_k^l\left|\mathbf{x}\right.\right)\exp\left\{\sum_{j\in\mathbb{J}_i}\phi_j^{-1}(\mathbf{x})\mathcal{P}_{j,k}^l\right\}} \tag{38}$$

with

$$\begin{aligned}\mathbb{X}_i^\varepsilon &= \left\{\mathbf{x}\in\mathcal{A}_1\times...\times\mathcal{A}_T\left|\phi_i^{-1}(\mathbf{x})=\varepsilon\right.\right\}\\ \mathbb{J}_i &= \left\{j\in 1,...,QT\left|j\neq i\right.\right\}\end{aligned} \tag{39}$$

B.2 List-APP sphere decoder

As a matter of fact, MAP MIMO detection will never be retained for high rate communication scenarii, since the point enumeration over the whole constellation subsets $\mathbb{X}_i^0$ and $\mathbb{X}_i^1$ rapidly becomes overwhelming for large numbers of transmit antennas and/or high-order modulations. A careful analysis of likelihood values reveals that a huge number of them are negligible. Hence, the point enumeration can be performed over much reduced size subsets $\mathbb{L}_i^1$ and $\mathbb{L}_i^0$, also called lists, which only contain the non-negligible likelihoods. In geometrical terms, those lists contain points of the lattice within a sphere centered on a well-chosen point (e.g., unconstrained ML or the ML point itself). This yields the approximation

$$\mathcal{L}_{i,k}^{l,\det} \approx \ln\frac{\sum_{\mathbf{x}\in\mathbb{L}_i^1} p\left(\mathbf{z}_k^l\left|\mathbf{x}\right.\right)\exp\left\{\sum_{j\in\mathbb{J}_i}\phi_j^{-1}(\mathbf{x})\mathcal{P}_{j,k}^l\right\}}{\sum_{\mathbf{x}\in\mathbb{L}_i^0} p\left(\mathbf{z}_k^l\left|\mathbf{x}\right.\right)\exp\left\{\sum_{j\in\mathbb{J}_i}\phi_j^{-1}(\mathbf{x})\mathcal{P}_{j,k}^l\right\}} \tag{40}$$

Lists may be efficiently exhibited using a modified version of the sphere decoder, whose detailed description can be found in [2].

B.3 IC-based detectors

To further reduce the complexity, joint MIMO detectors can be replaced by IC-based MIMO detectors, made of a bank of T linear filters and SISO detectors. Several kind of IC front-ends can be employed, namely SUMF and unconditional MMSE.

V. Monte-Carlo simulation results

In the present paper, our concern is to test the potential of the second approach in one of its simplest modes, i.e., MMSE MIMO ISI cancellation and SUMF-IC MIMO detection, and for modulation-coding schemes derived from [16]. We study the robustness of the turbo-equalizer to both ISI and MAI. The performance of the turbo-equalizer with MAP MIMO detection is given as a benchmark (when it can be simulated). Moreover, for all simulations, two analytical curves are systematically provided. The MF bound is used to validate the turbo-equalizer efficiency at fixed coding-modulation scheme[4]. The outage probability is plotted according to (7) as an absolute limit.

We start by investigating the robustness of the turbo-equalizer with respect to the channel selectivity. Following [16], we consider a STBICM using 8-PSK modulation (Gray labeling) and rate-1/3 64-state non-recursive convolutional (NRC) code. Code word length is $N = 1536$ bits (512 information bits), yielding $L = 256$ c.u per channel block. The MIMO channel is assumed block-static ($P = 1$) and comprises $T = 2$ transmit and $R = 2$ receive antennas. The spectral efficiency is $\eta = 2$ bits p.c.u. The channel memory is increased from $M = 1$ to $M = 9$ with equally (EQ) distributed fading matrix coefficients. In the turbo-equalizer, the order of the MMSE filter is set to $L_F = 9$ ($L_1 = L_2 = 4$) for $M = 1$ and $L_F = 21$ ($L_1 = L_2 = 10$) for $M = 9$. On Fig.1, we see how the choice of the detection algorithm can influence the convergence towards the MF bound when we suppose a weak channel dispersion. The simple SUMF-IC MIMO detection allows to approach the MF bound within 1.5 dB at BLER 10^{-2}. With MAP MIMO detection, the gap is reduced to 0.3 dB and the performance per iteration is always better. For a more severe channel dispersion (EQ-10 on Fig.2), the turbo-equalizer with SUMF-IC MIMO detection performs within 0.4 dB from the MF bound at BLER 10^{-2}. In that case, the ISI cancellation part works better, as predicted by the theory [9, theorem 1]. The second set of simulations (Fig.3,4) aims at testing the turbo-equalizer behavior when we increase the number of antennas at fixed dispersion (channel memory $M = 3$, EQ distributed fading matrix coefficients, filter order set to $L_F = 13$, $L_1 = L_2 = 6$). We keep 8-PSK modulation (Gray labeling) and rate-1/3 64-state NRC code. Code word length is $N = 1536$ bits for the block-static 4×4 system and $N = 3072$ bits for the block-static 8×8 system. The spectral efficiencies are $\eta = 4$ and 8 bits p.c.u, respectively. Again, the simplest combination of the proposed turbo-equalizer is sufficient to perform around 0.5 dB from the MF at targeted BLER 10^{-2}.

We conclude the section by few additional comments:

- In all transmission scenarii, STBICM-ID is able to perform quite close from the outage probability at BLER 10^{-2}. In the last three simulations, however, the slope of the MF bound differs from the slope of the outage probability, revealing that not all the promised diversity has been cap-

[4] The MF bound describes the hypothetical situation where perfect IC is performed.

tured. We will not go deeper into that interesting issue here.

- In [17] [18], it is demonstrated through Monte-Carlo simulations that if the channel is dispersive enough, the number of antennas sufficiently large, and the load per antenna not too heavy, simple matched filters may eventually replaced MMSE filters during the course of iterations without notable degradation in terms of performance. This observation is also valid for the second approach.
- This paper is focused on the data detection part of the receiver. Another iterative loop dealing with channel estimation (and reestimation) can easily complete the structure. An iterated-decision algorithm based on simple unconditional ML or MMSE criteria provides very good results [6].

VI. Conclusion

In this paper, the principle of iterated-decision linear equalization have been extended to STBICM over MIMO block fading multipath AWGN channel. Two different approaches have been described. Contrary to the first approach where ISI and MAI are jointly cancelled for each antenna, the second approach decouples the two tasks, allowing additional freedom degree in the receiver design. Simulation results (and underlying theoretical analysis) demonstrate the great potential of the second approach to realize powerful low-complexity equalizers with close to optimal performance. Emphasis was put on MMSE MIMO ISI cancellation and SUMF-IC MIMO detection, i.e., one of the simplest combinations which was proved to be sufficient for all presented transmission scenarii. However, for higher rates per transmit antenna, we have witnessed that the performance could severely degrade, entailing the resort to joint MIMO detection in MAP sense. The later can be efficiently performed by a list-APP sphere-decoder, whose complexity remains polynomial in all system parameters. This topic, as well as further research on the design of high-rate STBICM for MIMO block fading channel, will be addressed in future contributions.

References

[1] E. Biglieri, G. Tarrico, E. Viterbo, "Bit-Interleaved Time-Space Codes for Fading Channels," *Conf. Inform. Science and Systems*, Princeton University, New Jersey, Mar. 2000.

[2] B. M. Hochwald, S. Ten Brink, "Achieving Near-Capacity on a Multiple-Antenna Channel," *IEEE Trans. Commun.*, vol. COM-51, no. 3, pp. 389-399, May 2003.

[3] A.O. Berthet, R. Visoz, and J.J. Boutros, "Space-Time BICM versus Space-Time Trellis Code for MIMO Block Fading Multipath AWGN Channel," IEEE ITW'03, Paris, France, pp. 206-209, Mar./Apr. 2003.

[4] G. Bauch, H. Khorram, J. Hagenauer, "Iterative Equalization and Decoding in Mobile Communications Systems," EPMCC'97, Bonn, Germany, pp. 307-312, Sept. 1997.

[5] A. Tonello, "Space-Time Bit-Interleaved Coded Modulation over Frequency-Selective Fading Channels with Iterative Decoding," IEEE GLOBECOM'00, San Francisco, USA, pp. 1616-1620, 2000.

[6] R. Visoz, A.O. Berthet, "Iterative Decoding and Channel Estimation for Space-Time BICM over MIMO Block Fading Mutlipath AWGN Channel," *IEEE Trans. Commun.*, vol. COM-51, no. 8, pp. 1358-1367, Aug. 2003.

[7] A. Glavieux, C. Laot, J. Labat, "Turbo-Equalization over a Frequency Selective Channel," International Symposium on Turbo Codes, Brest, France, pp. 96-102, Sept. 1997.

[8] X. Wang, H.V. Poor, "Iterative (Turbo) Soft-Interference Cancellation and Decoding for Coded CDMA," *IEEE Trans. Commun.*, vol. COM-47, no. 7, pp. 1046-1061, July 1999.

[9] A.M. Chan, G.W. Wornell, "A Class of Block-Iterative Equalizers for InterSymbol Interference Channels: Fixed Channel Results," *IEEE Trans. Commun.*, vol. COM-49, no. 11, pp. 1966-1975, Nov. 2001.

[10] M. Tüchler, A.C. Singers, R. Koetter, "Minimum Mean Squared Error Equalization Using A Priori Information," *IEEE Trans. Signal Processing*, vol. SP-50, pp. 673-683, Mar. 2002.

[11] T. Abe, S. Tomisato, T. Matsumoto, "Performance Evaluation of a Sapce-Time Turbo-Equalizer in Frequency Selective MIMO Channels using Field Measurement Dara," IEE Seminar MIMO Commun. Systems, London, UK, pp. 21/1-21/5, Dec. 2001.

[12] T. Matsumoto, T. Abe, T. Yamada, "Towards Less than One Reuse Factor: Space-Time MIMO Turbo-Equalizers System-level Simulation Results," IEE Seminar MIMO Commun. Systems, London, UK, pp. 8/1-8/6, Dec. 2001.

[13] R. Knopp, "Coding and Multiple Access over Fading Channels," Doctoral Dissertation, EPFL, Lausanne, Switzerland, 1997.

[14] E. Biglieri, J. Proakis, S. Shamai (Shitz), "Fading Channels: Information-Theoretic and Communications Aspects," *IEEE Trans. Inform. Theory*, vol. IT-44, no. 6, pp. 2619-2692, Oct. 1998.

[15] L. Ozarow, S. Shamai (Shitz), A. Wyner, "Information Theoretic Considerations for Cellular Mobile Radio," *IEEE Trans. Vehic. Technol.*, vol. VT-43, no. 2, pp. 359-378, May 1994.

[16] S. Ariyavisitakul, " Turbo Space-Time Processing to Improve Wireless Channel Capacity," *IEEE Trans. Commun.*, vol. COM-48, no. 8, pp. 1347-1358, Aug. 2000.

[17] H. Oomori, T. Asai, T. Matsumoto, "A Matched-Filter Approximation for SC/MMSE Turbo-Equalizers," *IEEE Commun. Lett.*, vol. 5, no. 7, pp. 310-312, July 2001.

[18] K. Kansanen, T. Matsumoto, "A Computationally Efficient MIMO Turbo-Equalizer," *IEEE VTCS'03*, Seoul, Korea, pp. 277-281, Apr. 2003.

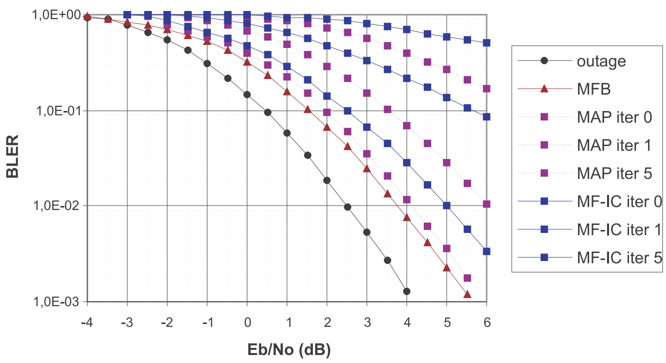

Fig. 1. Block-static MIMO system 2×2 EQ_2 η=2 bits p.c.u.

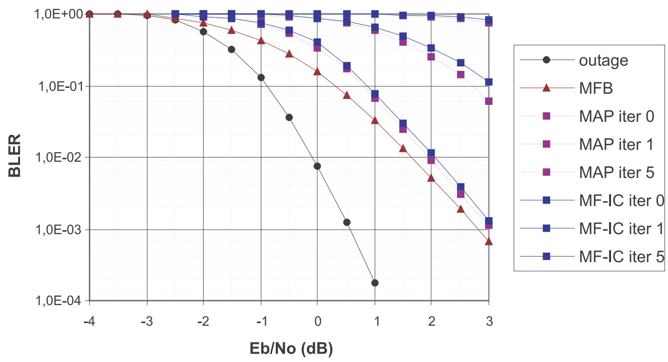

Fig. 2. Block-static MIMO system 2×2 EQ_{10} η=2 bits p.c.u.

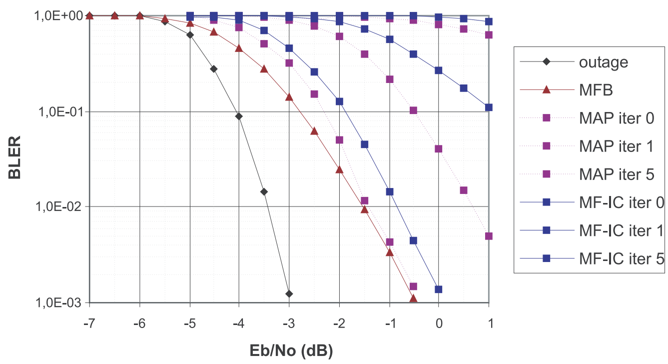

Fig. 3. Block-static MIMO system 4×4 EQ_4 η=4 bits p.c.u.

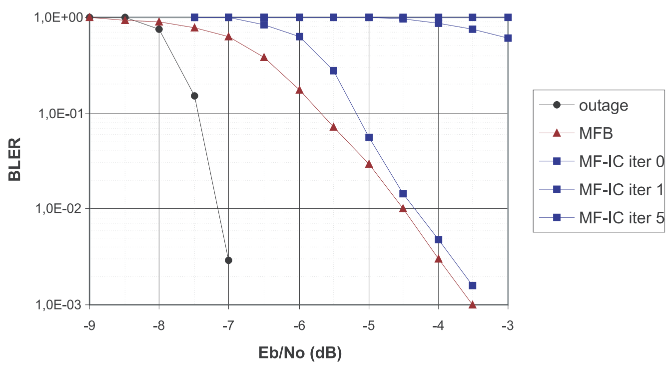

Fig. 4. Block-static MIMO system 8×8 EQ_4 η=8 bits p.c.u.