CHANNEL ESTIMATION IN TURBO-BLAST DETECTORS USING EM ALGORITHM

Mohammad-Ali Khalighi†, *Joseph Boutros*††

† Institut d'Électronique et de Télécommunications de Rennes (IETR), 20 av. des Buttes de Coësmes, 35043 Rennes, France
†† UMR 5141 CNRS - GET / Télécom Paris (ENST), 46 rue Barrault, 75013 Paris, France
Email: Mohammad-Ali.Khalighi@insa-rennes.fr, Boutros@comelec.enst.fr

ABSTRACT

We consider channel re-estimation based on EM (expectation maximization) algorithm in an iterative MIMO detector based on parallel interference cancellation, also known under the name of Turbo-BLAST. While providing an appropriate formulation of EM, we show that a primarily proposed EM implementation gives a biased estimate, and propose modifications in the EM implementation in order to obtain an unbiased estimate. Testing the proposed algorithms by simulations, we show that the achieved improvement is considerable for important diversity orders.

1. INTRODUCTION

Under suitable propagation conditions, multiple input multiple output (MIMO) systems can provide high spectral efficiencies. When coherent signal detection is used at receiver, a channel estimation task is unavoidable. This is usually based on some training (pilot) sequences, known to the receiver. When few pilot symbols are to be transmitted in each frame, use of semi-blind estimation techniques becomes of great interest.

In this paper, we consider the channel estimation in MIMO systems using iterative detection at receiver. The detection method we consider here is based on parallel interference cancellation (PIC), which is implemented in a turbo scheme together with a soft-input soft-output (SISO) decoder (Section 3). This detection scheme is also known under the name of Turbo-BLAST [1, 2]. Channel coefficients estimation update is performed in each iteration of the turbo-detector using the expectation maximization (EM) algorithm (Section 4). We first present the estimation based on pilot symbols only, and meanwhile, propose a modified estimation of the noise variance. We next consider a classical EM implementation, already proposed in [3], and propose an appropriate formulation for it under PIC. We next propose two modified methods for EM implementation to obtain an unbiased channel estimate. Performances of different estimation schemes are studied using simulations (Section 5). Frequency non-selective and uncorrelated quasi-static Rayleigh fading and QPSK modulation are considered throughout the paper.

2. SYSTEM MODEL

Consider a MIMO system with M_T and M_R antennas at transmitter and at receiver, respectively. We take $M_T = M_R$. Bit-interleaved coded modulation scheme is performed at transmitter using a non-recursive non-systematic convolutional (NRNSC) code (see Fig. 1). Encoded data bits are mapped to symbols x after being interleaved (block Π), and are then transmitted on M_T antennas. Symbols x are considered normalized in power and no precoding is performed on them prior to transmission. We will call a vector $\boldsymbol{x}$ of M_T symbols, a *compound symbol*.
Given the channel matrix $\boldsymbol{H}$ of dimension ($M_R \times M_T$), the vector of received signals at time sample n is

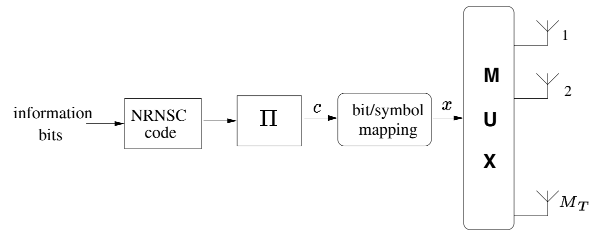

Fig. 1. Block diagram of the transmitter

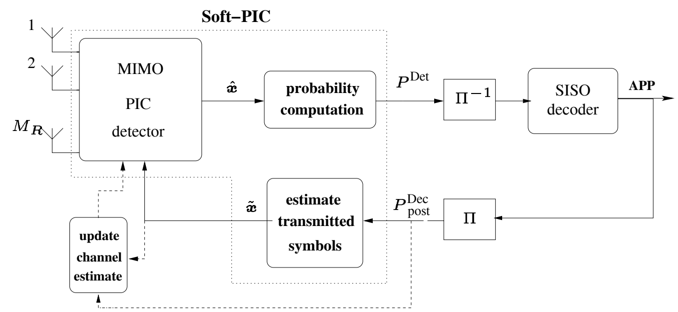

Fig. 2. Block diagram of Turbo-BLAST detector

$$\boldsymbol{y}(n) = \boldsymbol{H}\,\boldsymbol{x}(n) + \mathsf{n}(n) \tag{1}$$

n is the vector of complex circularly symmetric AWGN of variance σ^2.

3. TURBO-BLAST DETECTOR

The iterative detector is composed of two main blocks of *Soft-PIC* and SISO decoder, shown in Fig. 2. SISO decoder considered here is based on MAP algorithm. Soft-PIC is composed of three sub-blocks that we describe in the sequel.

• PIC detector
The PIC detection we consider here is based on minimum mean-square error (MMSE) detection and is described in detail in [1]. We just provide here the main detector expressions. Let $\boldsymbol{h}_k$ be the k^{th} column of $\boldsymbol{H}$. At the first iteration we dispose of no information on transmitted symbols. The detected k^{th} transmitted symbol (on the antenna $\#k$) at a given time sample is:

$$\hat{x}_k^{(1)} = \boldsymbol{h}_k^H \left(\boldsymbol{H}\boldsymbol{H}^H + \sigma^2 \boldsymbol{I}\right)^{-1} \boldsymbol{y} \tag{2}$$

The superscript $.^H$ denotes conjugate-transpose and $\boldsymbol{I}$ is the Identity matrix. From the second iteration, we can estimate the transmitted symbols $\tilde{\boldsymbol{x}}$ using the a posteriori probabilities at the SISO decoder output, and use them to remove (or better say, to reduce) the co-antenna interference (CAI). We use the following simplified expression for the detection of x_k and in this way avoid the inversion of an ($M_R \times M_R$) matrix in the exact formulation of PIC [1]. At the m^{th} iteration we have:

$$\hat{x}_k^{(m)} = \left(\boldsymbol{h}_k^H \boldsymbol{h}_k + \sigma^2\right)^{-1} \boldsymbol{h}_k^H \left(\boldsymbol{y} - \boldsymbol{H}_k\, \tilde{\boldsymbol{x}}_k^{(m-1)}\right) \tag{3}$$

where $\boldsymbol{H}_k$ of dimension $M_R \times (M_T - 1)$ is the matrix $\boldsymbol{H}$ with its k^{th} column removed and $\tilde{\boldsymbol{x}}_k^{(m-1)}$ of dimension $(M_T - 1) \times 1$ is the vector of estimated symbols at iteration $(m-1)$, with its k^{th} entry $\tilde{x}_k^{(m-1)}$ removed.

• Probability computation

Let for QPSK bit/symbol mapping, one bit $b_\Re$ corresponds to $\Re\{x\}$ and the other bit $b_\Im$ to $\Im\{x\}$ with $\Re\{.\}$ and $\Im\{.\}$ the real and imaginary part operators, respectively. Also, let $X_0 = -\sqrt{2}/2$ and $X_1 = -X_0$ be the symbol real part, respectively for bit=0 and bit=1 (the same for the symbol imaginary part). If the variance of noise + residual CAI (assumed of Gaussian distribution) equals σ'^2, the probability of $b_\Re = 1$, for instance, is then proportional to P_1, given below (σ'^2 can be estimated easily).

$$P_1 = \exp\left(-\frac{(\Re\{\hat{x}_k\} - \gamma_k X_1)^2}{\sigma'^2}\right) \quad ; \quad \gamma_k = \frac{\boldsymbol{h}_k^H \boldsymbol{h}_k}{\boldsymbol{h}_k^H \boldsymbol{h}_k + \sigma^2} \tag{4}$$

Similarly, the probability of $b_\Re = 0$, is proportional by P_0, obtained from (4) by replacing X_1 by X_0. Now,

$$\text{Prob}(b_\Re = 1) = \frac{P_1}{P_0 + P_1}. \tag{5}$$

Replacing $\Re$ by $\Im$ we obtain the corresponding expressions for the calculation of Prob$(b_\Im = 1)$.

• Estimation of transmitted symbols

Considering our convention on bit/symbol mapping we have:

$$\begin{array}{rcl} \Re\{\tilde{x}_k\} & = & APP_{k,r} \cdot X_1 + (1 - APP_{k,r}) \cdot X_0 \\ \Im\{\tilde{x}_k\} & = & APP_{k,i} \cdot X_1 + (1 - APP_{k,i}) \cdot X_0 \end{array} \tag{6}$$

where $APP_{k,r}$ and $APP_{k,i}$ denote the a posteriori probabilities corresponding to the real and imaginary parts of x_k, respectively.

4. CHANNEL ESTIMATION

4.1. Using pilot symbols only

Let $\boldsymbol{x}_p(n)$ denote a compound pilot symbol, i.e., a vector of M_T pilot symbols, at the time sample n. We define N_{ps} the number of compound pilot symbols in a frame of N_s compound symbols, that corresponds to $N_p = N_{ps} B M_T$ pilot bits, with B=2 the number of bits per symbol. Also, we define $N_{ds} = N_s - N_{ps}$, the number of compound data symbols in a frame. The ML channel estimate $\hat{\boldsymbol{H}}^p$ based on pilot symbols is:

$$\hat{\boldsymbol{H}}^p = \Big(\sum_{n=1}^{N_{ps}} \boldsymbol{y}(n)\, \boldsymbol{x}_p^H(n)\Big)\Big(\sum_{n=1}^{N_{ps}} \boldsymbol{x}_p(n)\, \boldsymbol{x}_p^H(n)\Big)^{-1} \tag{7}$$

To estimate the complex noise variance $\sigma^2 = N_0$ (required in (2), (3), and (4)), we may assume $\hat{\boldsymbol{H}}^p \approx \boldsymbol{H}$ and use the following solution:

$$\hat{N}_0 = \frac{1}{M_R N_{ps}}\Big(\sum_{n=1}^{N_{ps}} \|\boldsymbol{y}(n) - \hat{\boldsymbol{H}}^p\, \boldsymbol{x}_p(n)\|^2\Big) \tag{8}$$

However, this is a biased estimate of N_0, since $\hat{\boldsymbol{H}}^p$ contains estimation errors. Assuming mutually orthogonal pilot sequences, we propose in the following an unbiased estimation of N_0.

Let $\hat{\boldsymbol{H}}^p = \boldsymbol{H} + \boldsymbol{\Delta H}$ with $\boldsymbol{\Delta H}$ the matrix of estimation errors, and $\boldsymbol{\Delta y}(n) = \boldsymbol{y}(n) - \hat{\boldsymbol{y}}(n)$ with $\hat{\boldsymbol{y}}(n) = \hat{\boldsymbol{H}}^p \boldsymbol{x}_p(n)$. We have:

$$\boldsymbol{\Delta y}(n) = -\boldsymbol{\Delta H}\, \boldsymbol{x}_p(n) + \mathsf{n}(n) \tag{9}$$

where,

$$\boldsymbol{\Delta H} = \frac{1}{N_{ps}}\Big(\sum_{n=1}^{N_{ps}} \mathsf{n}(n)\, \boldsymbol{x}_p^H(n)\Big) \tag{10}$$

Equation (10) results from (7), assuming orthogonal training sequences satisfying $\boldsymbol{R}_{\boldsymbol{x}_p} = N_{ps} \boldsymbol{I}$. From (10) we obtain:

$$\begin{aligned} \boldsymbol{\Delta H}\, \boldsymbol{x}_p(n) &= \frac{1}{N_{ps}} \mathsf{n}(n)\, \|\boldsymbol{x}_p(n)\|^2 + \frac{1}{N_{ps}} \sum_{\substack{i=1 \\ i \neq n}}^{N_{ps}} \mathsf{n}(i)\, \boldsymbol{x}_p^H(i)\, \boldsymbol{x}_p(n) \\ &= \frac{M_T}{N_{ps}} \mathsf{n}(n) + \frac{1}{N_{ps}} \sum_{\substack{i=1 \\ i \neq n}}^{N_{ps}} \mathsf{n}(i)\, \boldsymbol{x}_p^H(i)\, \boldsymbol{x}_p(n) \end{aligned} \tag{11}$$

Now from (9) we have:

$$\boldsymbol{\Delta y}(n) = \Big(1 - \frac{M_T}{N_{ps}}\Big)\mathsf{n}(n) - \frac{1}{N_{ps}} \sum_{\substack{i=1 \\ i \neq n}}^{N_{ps}} \mathsf{n}(i)\, \boldsymbol{x}_p^H(i)\, \boldsymbol{x}_p(n) \tag{12}$$

Then,

$$\sum_{n=1}^{N_{ps}} \|\boldsymbol{\Delta y}(n)\|^2 \approx \Big(1 - \frac{M_T}{N_{ps}}\Big)^2 N_{ps} M_R\, N_0 \tag{13}$$

where we have assumed uncorrelated noise samples, large enough N_{ps}, and mutually orthogonal pilot sequences. If this third assumption is not exactly valid, we will have a relatively small positive offset in the obtained $\hat{N}_0$, which is quite tolerable. From (13) the new modified N_0 estimate is:

$$\hat{N}_0^{new} = \frac{N_{ps}}{M_R (N_{ps} - M_T)^2}\Big(\sum_{n=1}^{N_{ps}} \|\boldsymbol{y}(n) - \hat{\boldsymbol{H}}^p\, \boldsymbol{x}_p(n)\|^2\Big) \tag{14}$$

Notice that for $N_{ps} \gg M_T$, we find the primary estimate (8). However, for relatively small N_{ps}, the bias of (8) is important and the proposed modification is indispensable.

4.2. Classical EM-based channel estimation

Using the EM algorithm, while considering the ensemble of pilot and data symbols as *missing data*, we obtain the following estimation update equation for $\boldsymbol{H}$ at the iteration (i+1) [3]:

$$\hat{\boldsymbol{H}}^{(i+1)} = \overline{\boldsymbol{R}}_{yx}\, \overline{\boldsymbol{R}}_x^{-1} \tag{15}$$

$$\overline{\boldsymbol{R}}_{yx} = \sum_{n=1}^{N_s} \sum_{u=1}^{q} \boldsymbol{y}(n)\, \boldsymbol{x}_u^H\, APP_n(\boldsymbol{x}_u) \tag{16}$$

$$\overline{\boldsymbol{R}}_x = \sum_{n=1}^{N_s} \sum_{u=1}^{q} \boldsymbol{x}_u\, \boldsymbol{x}_u^H\, APP_n(\boldsymbol{x}_u) \tag{17}$$

$q = 2^{BM_T}$ is the cardinality of the constellation of compound symbols. Symbol $\boldsymbol{x}_u$ is the u^{th} among q possible compound symbols whose probability of transmission $APP_n(\boldsymbol{x}_u)$ is calculated using the a posteriori probabilities at the SISO decoder output:

$$APP_n(\boldsymbol{x}_u) \propto \prod_{i=1}^{BM_T} P_{\text{post}}^{\text{Dec}}(c_{u,i}) \tag{18}$$

$P_{\text{post}}^{\text{Dec}}(c_{u,i})$ is the a posteriori probability corresponding to the i^{th} bit of $\boldsymbol{x}_u$, $c_{u,i}$. For pilot bits, which are known at receiver, this equals either one or zero. At the first iteration, pilot sequences are used to obtain a primary channel estimate, and to permit the EM algorithm (used in the succeeding iterations) to *bootstrap*.

The computational complexity of (16) and (17) increases exponentially with BM_T. It can be shown that:

$$\sum_{u=1}^{q} \boldsymbol{x}_u APP(\boldsymbol{x}_u) = \tilde{\boldsymbol{x}} \tag{19}$$

$\tilde{\boldsymbol{x}}$ is already calculated in Soft-PIC and there is no need to recalculate it. Using (19), (16) and (17) can be written in the following form:

$$\overline{\boldsymbol{R}}_{yx} = \sum_{n=1}^{N_s} \boldsymbol{y}(n)\, \tilde{\boldsymbol{x}}^H(n) = \boldsymbol{H}\, \boldsymbol{R}'_x + \boldsymbol{\eta} \tag{20}$$

$$\overline{\boldsymbol{R}}_x(i,j) = \begin{cases} N_s & ; \quad i = j \\ \sum_{n=1}^{N_s} \tilde{\boldsymbol{x}}_i(n)\, \tilde{\boldsymbol{x}}_j^*(n) & ; \quad i \neq j \end{cases} \tag{21}$$

where $\tilde{\boldsymbol{x}}_i(n)$ is the i^{th} entry of the vector $\tilde{\boldsymbol{x}}(n)$. Also,

$$\boldsymbol{R}'_x \triangleq \sum_{n=1}^{N_s} \boldsymbol{x}(n)\, \tilde{\boldsymbol{x}}^H(n) \quad , \quad \boldsymbol{\eta} \triangleq \sum_{n=1}^{N_s} \mathsf{n}(n)\, \tilde{\boldsymbol{x}}^H(n) \tag{22}$$

$\boldsymbol{\eta}$ is the matrix of weighted noise samples with the auto-covariance matrix $N_0 \boldsymbol{R}''_x$, where,

$$\boldsymbol{R}''_x \triangleq \sum_{n=1}^{N_s} \tilde{\boldsymbol{x}}(n)\, \tilde{\boldsymbol{x}}^H(n) \tag{23}$$

In this way, EM is implemented in the body of Soft-PIC and in a much simpler way than by (16) and (17). The computational complexity of (20) and (21) now increases only linearly with BM_T.

4.3. Modifying Classical EM

From (15) and (20) the estimated channel matrix can be written as:

$$\hat{\boldsymbol{H}} = \boldsymbol{H}\, \boldsymbol{R}'_x\, \overline{\boldsymbol{R}}_x^{-1} + \boldsymbol{\eta}\, \overline{\boldsymbol{R}}_x^{-1} \tag{24}$$

We see from (21) and (22) that $\boldsymbol{R}'_x \neq \overline{\boldsymbol{R}}_x$, and hence, $\hat{\boldsymbol{H}}$ in (24) is a biased estimate. This comes from the fact that $\tilde{\boldsymbol{x}} \neq \boldsymbol{x}$ (except at high enough SNR where the bias becomes negligible).
In order to remove the bias we consider (separately) the pilot-only-based channel estimate $\hat{\boldsymbol{H}}^p$ and combine it with the data-based channel estimate (via EM) $\hat{\boldsymbol{H}}^d$ in an optimal manner to obtain $\hat{\boldsymbol{H}}$. In the sequel, we consider $\overline{\boldsymbol{R}}_x$, $\boldsymbol{R}'_x$, and $\boldsymbol{R}''_x$ given by (21), (22) and (23), calculated only over data symbols, that is, with the summations taken on N_{ds} instead of N_s.

4.3.1. Removing bias prior to combining

Let us define $\boldsymbol{B} = \boldsymbol{R}'_x\, \overline{\boldsymbol{R}}_x^{-1}$ the matrix of (multiplicative) bias on $\hat{\boldsymbol{H}}^d$. To remove this bias we can use the inverse filter $\boldsymbol{B}^{-1}$; however, the obtained results are not satisfying because of the resulting amplification of estimation errors. We use instead the MMSE filter $\boldsymbol{Q}$ given below:

$$\breve{\boldsymbol{H}}^d = \hat{\boldsymbol{H}}^d\, \boldsymbol{Q} \quad , \quad \boldsymbol{Q} = \left(\boldsymbol{B}^H \boldsymbol{B} + \boldsymbol{N}\right)^{-1} \boldsymbol{B}^H \; ; \tag{25}$$

$$\boldsymbol{N} = N_0\, \overline{\boldsymbol{R}}_x^{-H}\, \boldsymbol{R}''_x\, \overline{\boldsymbol{R}}_x^{-1} \tag{26}$$

So,

$$\breve{\boldsymbol{H}}^d = \boldsymbol{H}\boldsymbol{B}\left(\boldsymbol{B}^H \boldsymbol{B} + \boldsymbol{N}\right)^{-1} \boldsymbol{B}^H + \boldsymbol{\eta}'$$

We have $\boldsymbol{B}\left(\boldsymbol{B}^H \boldsymbol{B} + \boldsymbol{N}\right)^{-1} \boldsymbol{B}^H \approx \boldsymbol{I}$ and $\breve{\boldsymbol{H}}^d$ is almost unbiased.
Now we combine the entries of $\breve{\boldsymbol{H}}^d$ and $\hat{\boldsymbol{H}}^p$ so as to minimize the resulted corresponding estimation error variance:

$$\hat{\boldsymbol{H}}_{ij} = a_{ij}\, \breve{\boldsymbol{H}}_{ij}^d + b_{ij}\, \hat{\boldsymbol{H}}_{ij}^p \tag{27}$$

Let the variances of estimation errors in $\breve{\boldsymbol{H}}_{ij}^d$ and $\hat{\boldsymbol{H}}_{ij}^p$ be $\breve{\sigma}_{ij}^{d\,2}$ and $\hat{\sigma}_{ij}^{p\,2}$, respectively. Defining $\boldsymbol{Q}' = \overline{\boldsymbol{R}}_x^{-1}\, \boldsymbol{Q}$, it can be shown that:

$$\breve{\sigma}_{ij}^{d\,2} = N_0 \sum_{k=1}^{M_T} |\boldsymbol{Q}'_{kj}|^2\, \boldsymbol{R}''_{x\,kk} \quad ; \quad \hat{\sigma}_{ij}^{p\,2} = \frac{N_0}{N_{ps}} \tag{28}$$

It is seen that fortunately $\breve{\sigma}_{ij}^{d\,2}$ and $\hat{\sigma}_{ij}^{p\,2}$ do not depend on i, the index of the transmit antenna. So, the combination should be done equally within a column of the matrices:

$$\hat{\boldsymbol{H}}_j = a_j\, \breve{\boldsymbol{H}}_j^d + b_j\, \hat{\boldsymbol{H}}_j^p \tag{29}$$

and the optimization criterion is, hence,

$$\min \left\{ |a_j|^2 \breve{\sigma}_j^{d2} + |b_j|^2 \hat{\sigma}_j^{p2} \right\} \quad \text{subject to :} \quad a_j + b_j = 1. \tag{30}$$

Using the method of Largange multipliers, we obtain:

$$a_j = \frac{1}{1 + N_{ps} \sum_{k=1}^{M_T} |\boldsymbol{Q}'_{kj}|^2\, \boldsymbol{R}''_{x\,kk}} \quad , \quad b_j = 1 - a_j \tag{31}$$

In (26) we use the estimated N_0 at the previous iteration.

4.3.2. Compensating-bias combination

Here, before proceeding to the combination of estimates from data and pilot sequences, we consider some simplifying approximations. Inspiring by the idea of [4] for the case of CDMA multi-user detection, we assume that N_{ds} is large enough and approximate $\overline{\boldsymbol{R}}_x$, $\boldsymbol{R}'_x$, and $\boldsymbol{R}''_x$ by diagonal matrices. Indeed, the off-diagonal terms can be considered as empirical correlations between uncorrelated sequences and could be neglected. In this way, the (i,i)th entries of these matrices are:

$$\overline{\boldsymbol{R}}_{x\,ii} = N_{ds} \; , \; \boldsymbol{R}'_{x\,ii} = \sum_{n=1}^{N_{ds}} \boldsymbol{x}_i(n)\, \tilde{\boldsymbol{x}}_i^*(n) \; , \; \boldsymbol{R}''_{x\,ii} = \sum_{n=1}^{N_{ds}} |\tilde{\boldsymbol{x}}_i(n)|^2 \tag{32}$$

So, from (24) we have:

$$\hat{\boldsymbol{H}}^d = \frac{1}{N_{ds}} \boldsymbol{H}\; diag\left(\sum_{n=1}^{N_{ds}} \boldsymbol{x}_i(n) \tilde{\boldsymbol{x}}_i^*(n) \right) + \boldsymbol{\eta}'' \tag{33}$$

where $diag(.)$ denotes a diagonal matrix with its (i,i)th entry given, and $\boldsymbol{\eta}'' = \frac{1}{N_{ds}} \boldsymbol{\eta}$. The (i,j)th entry of $\hat{\boldsymbol{H}}^d$ is then,

$$\hat{\boldsymbol{H}}_{ij}^d = \alpha_j \boldsymbol{H}_{ij} + \boldsymbol{\eta}''_{ij} \tag{34}$$

where,

$$\alpha_j \triangleq \frac{1}{N_{ds}} \sum_{n=1}^{N_{ds}} \boldsymbol{x}_j(n)\, \tilde{\boldsymbol{x}}_j^*(n) \tag{35}$$

Now, thanks to the simplifying assumption of diagonality of $\boldsymbol{B}$, we can directly combine $\hat{\boldsymbol{H}}^d$ and $\hat{\boldsymbol{H}}^p$ and compensate the bias meanwhile, without any need to performing a pre-filtering on $\hat{\boldsymbol{H}}^d$. Notice that this diagonality assumption (particularly for $\boldsymbol{R}'_x$) has the meaning of perfect CAI cancellation by PIC, which is not really true, especially for low SNR and at first iterations of the detector. Anyway, the optimization problem has now become:

$$\hat{\boldsymbol{H}}_j = a_j \hat{\boldsymbol{H}}_j^d + b_j \hat{\boldsymbol{H}}_j^p \tag{36}$$

$$\min \left\{ |a_j|^2 \hat{\sigma}_j^{d2} + |b_j|^2 \hat{\sigma}_j^{p2} \right\} \quad \text{subject to :} \quad a_j \alpha_j + b_j = 1 \tag{37}$$

$\hat{\sigma}_j^{d\,2}$ is the variance of $\boldsymbol{\eta}''_j$ entries, which equals β_j^2 / N_{ds}, where,

$$\beta_j^2 = \boldsymbol{R}''_{x\,jj} = \frac{1}{N_{ds}} \sum_{n=1}^{N_{ds}} |\tilde{\boldsymbol{x}}_j(n)|^2. \tag{38}$$

The optimal combination coefficients are given by:

$$a_j = \frac{\alpha_j^*}{|\alpha_j|^2 + \frac{N_{ps}}{N_{ds}}\, \beta_j^2} \quad , \quad b_j = \frac{\frac{N_{ps}}{N_{ds}}\, \beta_j^2}{|\alpha_j|^2 + \frac{N_{ps}}{N_{ds}}\, \beta_j^2} \tag{39}$$

To calculate α_j, we notice that for example, $\Re\{\boldsymbol{x}_j(n)\} = \Re\{\tilde{\boldsymbol{x}}_j(n)\}$ with probability $(1-\varepsilon)$, where ε is the error probability on coded bits.

Considering our convention on bit/symbol QPSK mapping, it can be shown that:

$$\alpha_j \; N_{ds} \approx (1-2\varepsilon) X_1 \sum^{N_{ds}} \bigg(\Big[\big|\Re\{\tilde{\boldsymbol{x}}_j\}\big| + \big|\Im\{\tilde{\boldsymbol{x}}_j\}\big| \Big]$$

$$+\sqrt{-1}\Big[\Re\{\tilde{\boldsymbol{x}}_j\}\cdot \mathrm{sgn}\big(\Im\{\tilde{\boldsymbol{x}}_j\}\big) - \Im\{\tilde{\boldsymbol{x}}_j\}\cdot \mathrm{sgn}\big(\Re\{\tilde{\boldsymbol{x}}_j\}\big)\Big]\bigg) \quad (40)$$

where $\mathrm{sgn}(.)$ is the sign function. However, practically, we can take $\varepsilon = 0$ (which is true at high SNR) without observing any considerable change in the performance.

• *Interpreting the proposed optimal combinations*

In low SNR and at first iterations, the decoder output APPs are not reliable enough, and hence, logically we should give much more importance to $\hat{\boldsymbol{H}}^p$. This is verified in both combination solutions where for $\varepsilon \to 0.5$ we have $a_j \to 0$ and $b_j \to 1$. On the other hand, in high SNR and at concluding iterations, the estimates based on pilot and data symbols have almost the same reliability and should be combined equally. Both solutions verify this by $a_j \to N_{ps}/(N_{ps}+N_{ds})$ and $b_j \to N_{ds}/(N_{ps}+N_{ds})$.

5. SIMULATION RESULTS AND CONCLUSIONS

We compare the performances of different estimation methods for different system parameters. Channel code is an NRNSC of constraint length K=3 and interleaver is of random type. Performance is measured in terms of frame error rate (FER). We have shown in Figures 3, 4, and 5 curves of FER versus average E_b/N_0 at receiver for 4×4, 8×8, and 10×10 systems. Parameters N_p and N_d indicated on the figures represent respectively the number of pilot and data bits in a frame. Almost the same N_{ds} is considered, since we compare FERs. The number N_{ps} is chosen a little larger than the limit of identifiability BM_T. FER curves are shown for perfect channel knowledge and estimations based on: pilots only, classical EM (EM-Mix), removed bias EM by filtering $\hat{\boldsymbol{H}}^d$ (RB-EM), and modified unbiased EM by approximate combination (MU-EM). For comparison, we have also shown the outage probability curves.

We see that the performance obtained by RB-EM is quite comparable to that by MU-EM. Notice that even MMSE filtering of (25) tries to inverse "partially" the matrix $\boldsymbol{B}$, and hence, increases the estimation error variance. So, although RB-EM do not rely on approximations as it is the case for MU-EM, its performance is not considerably better. So, MU-EM, which is simpler to implement, is preferred. On the other hand, MU-EM performs better than EM-Mix only when the assumption of matrices diagonality made in Subsection 4.3.2 is closely satisfied. This returns to the assumption of perfect CAI cancellation, which for PIC detection holds when enough receive diversity is available.We see from simulation results that it is the case for 8×8 and 10×10 systems. The interesting result is that this is not the case for the 4×4 system, although the CAI comes from a smaller number of antennas. In other words, for larger number of antennas, the increased diversity gain overtakes the increased CAI, and overall, PIC performs better.

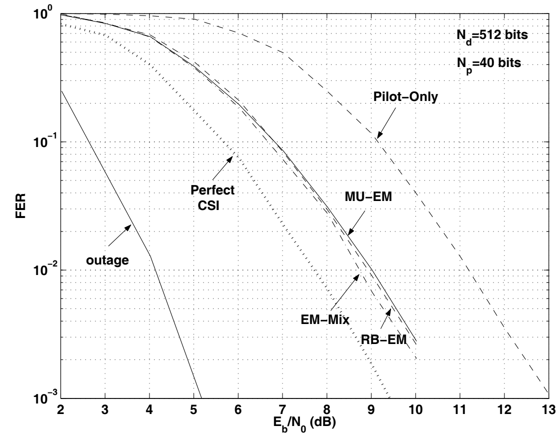

Fig. 3. M_T=M_R=4, NRNSC code [5,7], $N_p = 40$

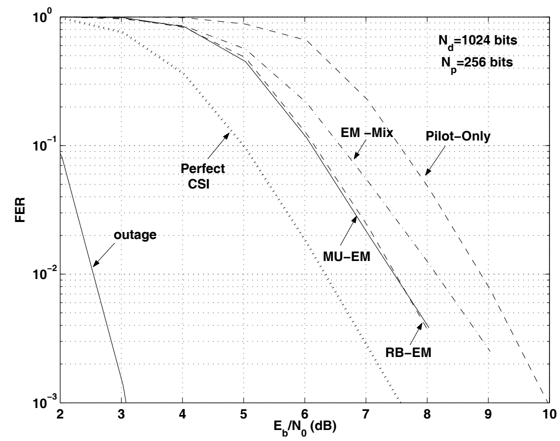

Fig. 4. M_T=M_R=8, NRNSC code [5,7], $N_p = 256$

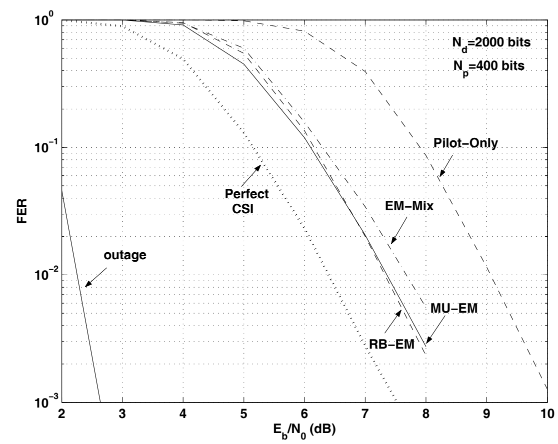

Fig. 5. M_T=M_R=10, NRNSC code [5,7], $N_p = 400$

6. REFERENCES

[1] M. Sellathurai and S. Haykin, "Turbo-BLAST for wireless communications: theory and experiments," *IEEE Transactions on Signal Processing*, vol. 50, no. 10, pp. 2538– 2546, Oct. 2002.

[2] G.D. Golden, G.J. Foschini, R.A. Valenzuela, and P.W. Wolniansky, "Detection algorithm and initial laboratory results using V-BLAST space-time communication architecture," *Electronic Letters*, vol. 35, no. 1, pp. 14–16, Jan. 1999.

[3] J. Boutros, F. Boixadera, and C. Lamy, "Bit-interleaved coded modulations for multiple-input multiple-output channels," *IEEE International Symposium on Spread Spectrum Techniques and Applications*, vol. 1, pp. 123 – 126, Sept. 2000.

[4] M. Kobayashi, J. Boutros, and G. Caire, "Successive interference cancellation with SISO decoding and EM channel estimation," *IEEE Journal on Selected Areas in Communications*, vol. 19, no. 8, pp. 1425–1428, Aug. 2001.