# Joint channel estimation and simplicity-based detection for large-scale MIMO FEC-coded systems

Zahran Hajji, Karine Amis Cavalec, Abdeldjalil Aissa El Bey

# Joint channel estimation and simplicity-based detection for large-scale MIMO FEC-coded systems

Zahran Hajji     Karine Amis     Abdeldjalil Aïssa-El-Bey
IMT Atlantique, UMR CNRS 6285 Lab-STICC, UBL, F-29238 Brest, France
Email: Firstname.Lastname@imt-atlantique.fr

***Abstract*—In this paper, we address the problem of channel estimation and signal detection in large MIMO FEC-coded systems assuming finite alphabet modulations. We consider a semi-blind iterative expectation maximization algorithm which relies on a limited number of pilot sequences to initialize the estimation process. We propose to include the estimation process within a turbo finite-alphabet simplicity (FAS)-based detection receiver. To that purpose we define two estimation updates from the FEC decoder output. Simulations carried out in both determined and undetermined configurations show that the resulting scheme outperforms the state-of-the-art receiver which uses an MMSE estimation criterion and that it reaches the maximum-likelihood lower-bound.**

***Index Terms*—Channel estimation, simplicity, signal detection, massive MIMO, turbo-detection**

## I. Introduction

Massive MIMO is considered as a potential candidate to address the challenges of 5G. The idea is to implement a large number of antennas to better exploit the spatial diversity so as to provide higher throughput under spectrum limitations. Efficient channel estimation with a minimum overhead is a key step to take up the challenge.

A first family of channel estimation techniques, called training-based estimation, uses pilot sequences inserted in the transmitted frame and perfectly known at the receiver [1]. In the case of massive MIMO which involves a high number of channel impulse responses, the definition of orthogonal sequences is an issue and the required training sequence length and number may become huge resulting in a waste of bandwidth and power. To overcome the pilot sequences issue, a second class of techniques referred to as blind estimation exploits received data only [2]. Unfortunately, in this case, channel parameters may be identified within scaling and permutation ambiguities which limits their exploitation.

To meet a trade-off between waste of bandwidth and power on one hand and, blind channel estimation ambiguities on the other hand, an intermediate solution is semi-blind estimation [3]: it uses a limited number of pilot sequences compared to training-based methods and improves the quality of channel estimation compared to blind estimation algorithms [4]. The idea here is to get a first raw estimate of the channel using pilot symbols and then, to progressively improve its accuracy by considering recovered symbols to define extended pilot sequences within an iterative estimation procedure [4], [5]. Following this approach, one can either improve channel estimation or reduce the required pilot symbol number.

The authors would like to thank the Brittany Region and PRACOM for their financial support.

Different ways to refine the channel state information (CSI) after a first raw training-based estimation exist as for example Kalman estimators [6]. The so-called Expectation-Maximization (EM) algorithm is particularly interesting. It provides a framework to iteratively calculate probabilities of unknown data elements to be used to estimate channel parameters.

Decision-directed estimators are fed with hard decisions taken on detection output. The risk is poor performance due to information loss especially with MMSE-based detection [7]. Substituting hard decisions with soft decisions provides confidence measures which iteratively improves channel estimation and makes it more accurate [8].
An outer FEC is usually applied before the modulation. Turbo-like receivers based on iterative information exchanges between their components (detection, decoder, channel estimation, synchronisation, ...) [9] were proved efficient to achieve near-optimal performance. To that purpose, the authors of [10] proposed to associate an EM-based detection and estimation with a SISO decoder within an iterative process.

In this paper, we consider large-scale MIMO FEC-coded systems. We address the problem of channel estimation and signal detection. We first describe the turbo finite-alphabet simplicity-based (FAS) detection scheme introduced in [11] for a perfect channel knowledge. Then we consider the iterative channel estimation method defined and analysed in [12] which combines FAS detection with EM algorithm. Our goal is the definition of a turbo FAS-based receiver which iteratively and successively performs large-scale MIMO channel estimation, data detection and FEC decoding.

Our contributions are: ***(i)*** a turbo FAS-based detection receiver which combines estimation, detection and FEC decoding. ***(ii)*** two ways of updating the FAS-based EM channel estimation from FEC decoder output.

This paper is organized as follows. Section II describes the large-scale MIMO system model. Section III details the turbo FAS-detection scheme introduced in [11] for perfect channel knowledge. Section IV deals with the introduction of the FAS-based EM channel estimation within the turbo receiver. We propose two ways of exploitation of FEC decoder output to update the EM channel estimation. Section V is dedicated to simulation results. Finally, Section VI concludes the paper.

**Notations:** boldface upper case letters and boldface lower case letters denote matrices and vectors, respectively. For the transpose, transpose conjugate and conjugate matrices we use $(\cdot)^T$,$(\cdot)^H$ and $(\cdot)^*$, respectively. $\otimes$ is the Kronecker product. $\boldsymbol{I}_p$ is the $p \times p$ identity matrix and $\mathbf{1}_p$ is the all-one size-$p$ vector. Given a complex-valued vector $\boldsymbol{x}$, its real-valued transform is the vector $\underline{\boldsymbol{x}}$ defined by $\underline{\boldsymbol{x}} = \begin{bmatrix}\Re(\boldsymbol{x}) & \Im(\boldsymbol{x})\end{bmatrix}^T$. Given a complex-valued matrix $\boldsymbol{X}$, its real-valued transform is the matrix $\underline{\boldsymbol{X}} = \begin{pmatrix}\Re(\boldsymbol{X}) & -\Im(\boldsymbol{X}) \\ \Im(\boldsymbol{X}) & \Re(\boldsymbol{X})\end{pmatrix}$.

## II. SYSTEM MODEL

Let us consider a large-scale MIMO system equipped with $N$ antennas at the transmitter and $n$ antennas at the receiver ($n \leq N$). Each transmitted frame consists of $T_p$ pilot vectors and $T_d$ data vectors ($T = T_p + T_d$). Under the above assumptions, the received signal can be modelled as:

$$\boldsymbol{Y} = \boldsymbol{H}\boldsymbol{X} + \boldsymbol{Z}. \tag{1}$$

$\boldsymbol{Y} = (\boldsymbol{Y}_p, \boldsymbol{Y}_d)$ is the received signal matrix. $\boldsymbol{Y}_p$ and $\boldsymbol{Y}_d$ are the $n \times T_p$ pilot received matrix and the $n \times T_d$ data received matrix respectively. $\boldsymbol{X}$ stands for the transmitted signal matrix. This $N \times T$ complex matrix can be decomposed as $\boldsymbol{X} = (\boldsymbol{X}_p, \boldsymbol{X}_d)$. $\boldsymbol{X}_p$ and $\boldsymbol{X}_d$ are the $N \times T_p$ pilot transmitted matrix and the $N \times T_d$ data transmitted matrix respectively.

We denote by $\boldsymbol{x}$ the $t$-th column of $\boldsymbol{X}$ which is the transmitted vector at time $t$. Its $k$-th element $x_k$ belongs to $\mathcal{B} = \{\beta_1, \beta_2, .., \beta_M\}$ such that its real and imaginary parts take values on $\mathcal{F} = \{\alpha_1, \alpha_2, .., \alpha_p\}$ where $p = \sqrt{M}$. Pilot symbols are assumed to be known at the BS and to be mutually orthogonal: $\boldsymbol{x}\boldsymbol{x}^H = \boldsymbol{I}_N$ for all $\boldsymbol{x}$ with $t = 0, \ldots, T_p - 1$. Transmitted data symbols are independent and identically distributed (i.i.d.): $\mathbb{E}[\boldsymbol{x}\boldsymbol{x}^H] = \boldsymbol{I}_N$ for all $\boldsymbol{x}$ with $t = T_p, \ldots, T$.

$\boldsymbol{Z}$ is the $n \times T$ noise matrix. Its components are i.i.d.complex circularly-symmetric complex Gaussian variables with zero mean and variance $\sigma^2$. $\boldsymbol{H}$ is an $n \times N$ complex random matrix which stands for the channel matrix. It can be written as

$$\boldsymbol{H} = \boldsymbol{G}\boldsymbol{\Gamma}^{\frac{1}{2}}, \tag{2}$$

where $\boldsymbol{G}$ is an $n \times N$ matrix representing small-scale fading with *i.i.d.* coefficients and $\boldsymbol{\Gamma}$ is a diagonal matrix which models large-scale fading. Its coefficients account for the path loss and shadow fading. We assume that the the columns of $\boldsymbol{G}$ are independent from $\boldsymbol{\Gamma}$ and are *i.i.d* circularly-symmetric complex normal vectors. We consider a block fading model: channel coefficients are constant over a frame of $T$ symbols and change independently at next coherence block (time division duplex assumption). The uplink received signal at time $t$ is denoted by $\boldsymbol{y}$ and is given by

$$\boldsymbol{y} = \boldsymbol{H}\boldsymbol{x} + \boldsymbol{z}, \tag{3}$$

We first transform the complex-valued system into an equivalent real-valued system, which reads

$$\underline{\boldsymbol{y}} = \underline{\boldsymbol{H}}\underline{\boldsymbol{x}} + \underline{\boldsymbol{z}}, \tag{4}$$

where $\underline{\boldsymbol{y}}$, $\underline{\boldsymbol{x}}$, $\underline{\boldsymbol{z}}$ and $\underline{\boldsymbol{H}}$ are the real-valued transforms of $\boldsymbol{y}$, $\boldsymbol{x}$, $\boldsymbol{z}$ and $\boldsymbol{H}$ respectively (cf. notations paragraph). The detection described hereinafter applies on this model.

## III. TURBO DETECTION ASSUMING PERFECT CHANNEL ESTIMATION

In this section, we consider that the binary stream is FEC encoded, then randomly interleaved before being converted into QAM symbols and passed through a serial-to-parallel converter. We briefly remind the turbo receiver based on the FAS detection and proposed in [11].

*1) FAS detection:* Let us describe the first detection iteration which corresponds to the FAS algorithm introduced in [13]. The data vector $\underline{\boldsymbol{x}}$, for all $t$-th column with $t = T_p + 1, \ldots, T$, is simple as its components belong to the interval $[\alpha_1, \alpha_p]$. It can be decomposed as $\underline{\boldsymbol{x}} = \boldsymbol{B}_\alpha \boldsymbol{r}$ where $\boldsymbol{B}_\alpha = \boldsymbol{I}_{2N} \otimes [\alpha_1, \alpha_p]$ and $\boldsymbol{r} \in [0,1]^{4N}$. Based on this decomposition, we proposed to solve the following optimization problem [13]:

$$\begin{aligned} \arg\min_{\boldsymbol{r}} \|\underline{\boldsymbol{y}} - \underline{\boldsymbol{H}}\boldsymbol{B}_\alpha \boldsymbol{r}\|_2 \quad \text{subject to} \\ \boldsymbol{B}_1 \boldsymbol{r} = \mathbf{1}_{2N} \quad \text{and} \quad \boldsymbol{r} \geq 0, \end{aligned} \tag{5}$$

where $\boldsymbol{B}_1 = \boldsymbol{I}_{2N} \otimes \mathbf{1}_2^T$. The problem (5) can be solved by the simplex [13] or the interior point methods [14]. In this paper, we consider interior point methods. These algorithms start by finding an interior point of the polytope defined by the constraints and then proceed to the optimal solution by moving inside the polytope.

The detector provides the FEC decoder with interleaved log-likelihood ratios which are denoted by $\Lambda_{in}^{dec}$ and whose definition requires the statistics of the detector output. Let $\hat{\boldsymbol{r}}$ stand for the solution of (5). The components of $\hat{\underline{\boldsymbol{x}}} = \boldsymbol{B}_\alpha \hat{\boldsymbol{r}}$ follow a censored normal distribution (combination of binary distributions on the bounds and Gaussian ones inside), which is given by [13] [11]:

$$f_{\hat{\underline{x}}_k}(x) = \frac{1}{p} \sum_{\ell=1}^{p} f_{\hat{\underline{x}}_k | x_k = \alpha_\ell}(x), \tag{6}$$

with

$$\begin{aligned} f_{\hat{\underline{x}}_k | \boldsymbol{x}_k = \alpha_\ell}(x) = \Bigg( & \frac{1}{2}\text{erfc}\left(\frac{\alpha_\ell - \alpha_1}{\sqrt{2}\sigma_{\hat{\underline{x}}}}\right)\delta_{\alpha_1}(x) \\ & + \frac{1}{2}\text{erfc}\left(\frac{\alpha_p - \alpha_\ell}{\sqrt{2}\sigma_{\hat{\underline{x}}}}\right)\delta_{\alpha_p}(x) \\ & + \frac{1}{\sqrt{2\pi}\sigma_{\hat{\underline{x}}}} \exp\left(-\frac{(x - \alpha_\ell)^2}{2\sigma_{\hat{\underline{x}}}^2}\right) \mathbb{1}_{[\alpha_1, \alpha_p]}(x) \Bigg). \end{aligned} \tag{7}$$

and

$$\sigma_{\hat{\underline{x}}}^2 = \sum_{k=0}^{2n-2} \binom{2N}{k} \left(\frac{1}{p}\right)^{2N-k} \left(\frac{p-1}{p}\right)^k \frac{2n\sigma^2}{2n-k-1}, \tag{8}$$

where $\delta_\alpha(x)$ is the Dirac delta function concentrated at $\alpha$ and $\mathbb{1}_\Omega(x)$ is the indicator function of the subset $\Omega$.

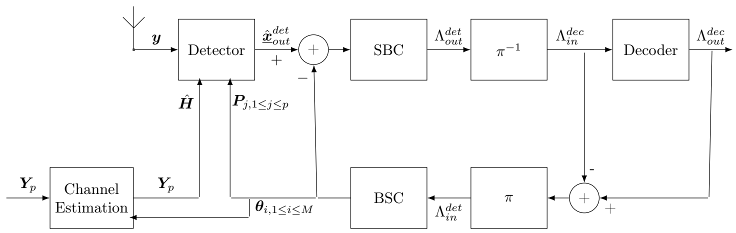


Fig. 1: Turbo detection scheme

Let us now describe next detection iterations [11]. The detector is fed with log-likelihood ratios denoted by $\Lambda_{in}^{det}$ and computed from the FEC decoder output. The probability vector $\boldsymbol{P}_j$ defined from $\Lambda_{in}^{det}$ by

$$\boldsymbol{P}_j = \left[\Pr(\underline{x}_1 = \alpha_j|\Lambda_{in}^{det}), \ldots, Pr(\underline{x}_{2N} = \alpha_j|\Lambda_{in}^{det})\right]^T, \quad (9)$$

is used to compute and define the Mean Absolute Error (MAE) as a regularization term in the following optimization problem [11]:

$$\underset{\boldsymbol{B}_1\boldsymbol{r}=\boldsymbol{1}_{2N},\boldsymbol{r}\geq 0.}{\arg\min} ||\underline{\boldsymbol{y}} - \underline{\boldsymbol{H}}\boldsymbol{B}_\alpha\boldsymbol{r}||_2 + \gamma\sum_{j=1}^{p}\boldsymbol{P}_j^T\,|\boldsymbol{r}-\boldsymbol{d}_j|, \quad (10)$$

with $\boldsymbol{d}_j = \alpha_j \times \boldsymbol{1}_{2N}$ and $\gamma = \sigma\sqrt{\frac{\log N}{n}}$.

On one hand, the regularization term can be seen as a penalty, imposed to ensure that the detector output remains in the neighborhood of the decoder output. On the other hand, $\gamma$ enables to regulate the contribution of the FEC information and thereby to question the FEC decision if necessary.

*2) Symbol to Binary Converter:* Let $m = \log_2(p)$ and $\boldsymbol{c}$ be the length-$2mN$ coded and interleaved binary information sequence at one channel use. Let also $\psi$ be the binary-to-symbol conversion defined as:

$$\psi : [c_{km}\ c_{km+1}\ \ldots\ c_{(k+1)m-1}] \in \{0,1\}^m \mapsto \underline{x}_k \in \mathcal{F} \quad (11)$$

and $\boldsymbol{c}^{(j)} = \psi^{-1}(\alpha_j)$.

Let $\hat{\underline{\boldsymbol{x}}}_{out}^{det}$ stand for the detector output. The symbol-to-binary converter (SBC) computes the log likelihood ratio on the $i$-th bit associated to the $k$-th symbol, denoted by $\Lambda_{out}^{det}$ and defined by:

$$\Lambda_{out}^{det}(km+i) = \log\left(\frac{Pr(c_{km+i}=1|y)}{Pr(c_{km+i}=0|y)}\right) \quad (12)$$

$$= \log\left(\frac{\sum_{\alpha_j\in\mathcal{F}_{i,1}} f_{\hat{\underline{x}}_k|\underline{x}_k=\alpha_j}(\hat{\underline{x}}_{out,k}^{det})Pr(\underline{x}_k=\alpha_j|\Lambda_{in}^{det}))}{\sum_{\alpha_j\in\mathcal{F}_{i,0}} f_{\hat{\underline{x}}_k|\underline{x}_k=\alpha_j}(\hat{\underline{x}}_{out,k}^{det})Pr(\underline{x}_k=\alpha_j|\Lambda_{in}^{det}))}\right)$$

with $\mathcal{F}_{i,\epsilon} = \{a \in \mathcal{F}|\boldsymbol{c} = \psi^{-1}(a), c_i = \epsilon\}$.

Let us mention that an empirical study proved that the expression of $\sigma_{\hat{\underline{x}}}$ given by (8) keeps valid throughout the iterative process.

## IV. Channel estimation algorithm

### A. Overview of EM estimator

The Maximum Likelihood (ML) estimate of $\boldsymbol{H}$ is given by

$$\hat{\boldsymbol{H}}_{ML} = \arg\max_{\boldsymbol{H}} \log p(\boldsymbol{Y}|\boldsymbol{H}) \quad (13)$$

As data symbols are not known, the ML problem cannot be analytically solved. It is necessary to use iterative algorithms that converge to the solution of (13). Among them, the EM algorithm updates the channel estimate based on an old one in the following manner:

$$\hat{\boldsymbol{H}}_{i+1} = \arg\max_{\boldsymbol{H}} \mathbb{E}[p(\boldsymbol{X}_d|\boldsymbol{Y},\hat{\boldsymbol{H}}_i)](\log p(\boldsymbol{Y},\boldsymbol{X}_d|\boldsymbol{H})). \quad (14)$$

As we can see, the algorithm involves an expectation step and a maximization one. The maximization step can be simplified and the updated estimate of the channel matrix can be written as

$$\hat{\boldsymbol{H}}_{i+1} = (\boldsymbol{Y}_p\boldsymbol{X}_p^H + \boldsymbol{Y}_d\,\mathbb{E}[\boldsymbol{X}_d|\boldsymbol{Y},\hat{\boldsymbol{H}}_i]^H) \quad (15)$$
$$\times(\boldsymbol{X}_p\boldsymbol{X}_p^H + \mathbb{E}[\boldsymbol{X}_d\boldsymbol{X}_d^H|\boldsymbol{Y},\hat{\boldsymbol{H}}_i])^{-1}.$$

### B. EM channel estimation algorithm and FAS-detection

Contrary to [12] which considered uncoded systems, we propose to feed the channel estimation with the FEC decoder output. The proposed semi-blind iterative EM algorithm steps are described hereinafter.

*1) Initialization step:* The first iteration only uses pilot sequences and relies on a maximum likelihood channel estimate, which is given by

$$\hat{\boldsymbol{H}}_0 = (\boldsymbol{Y}_p\boldsymbol{X}_p^H)(\boldsymbol{X}_p\boldsymbol{X}_p^H)^{-1}. \quad (16)$$

The corresponding mean square error (MSE) is computed as

$$\mathbb{E}(||\boldsymbol{H} - \hat{\boldsymbol{H}}_0||_2^2) = \frac{nN\sigma^2}{T_p} \quad (17)$$

Then, the receiver handles the following model:

$$\boldsymbol{y} \;=\; \hat{\boldsymbol{H}}_0\,\boldsymbol{x} + (\boldsymbol{H} - \hat{\boldsymbol{H}}_0)\,\boldsymbol{x} + \boldsymbol{z} = \hat{\boldsymbol{H}}_0\,\boldsymbol{x} + \boldsymbol{\eta},$$

where $\boldsymbol{\eta}$ is the updated additive Gaussian noise vector with zero mean and covariance matrix $(\frac{nN}{T_p}+1)\sigma^2 I_N$. The variance of real-valued FAS-detected vector is then calculated as follows:

$$\sigma_{\hat{\underline{x}}}^2 = \sum_{k=0}^{2n-2}\binom{2N}{k}\left(\frac{1}{p}\right)^{2N-k}\left(\frac{p-1}{p}\right)^k\frac{2n(\frac{nN}{T_p}+1)\sigma^2}{2n-k-1}. \quad (18)$$

To update the channel estimation, we propose two approaches that exploit the probability vectors $\boldsymbol{P}_j$ delivered by the decoder: the first based on hard-decisions and the second on soft-decisions.

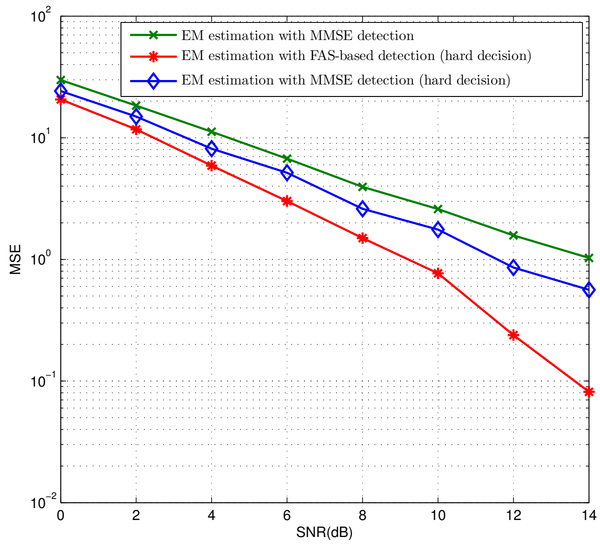


Fig. 2: MSE versus SNR with uncoded 4-QAM, $n = N = 64$, $T_p = 160$ and $T = 1024$.

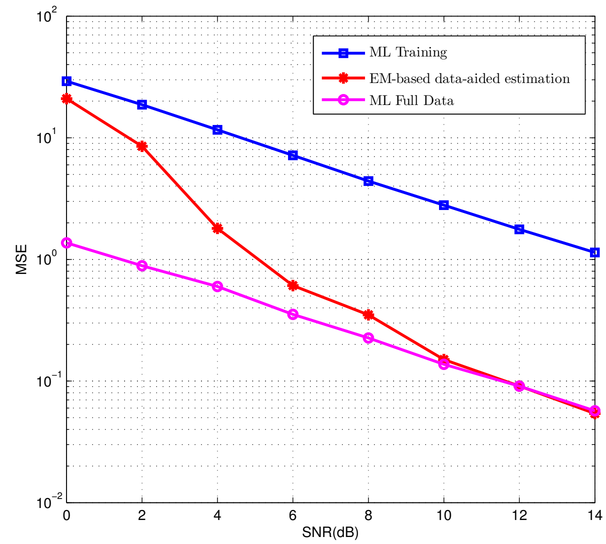


Fig. 3: MSE versus SNR with coded 4-QAM, $n = N = 64$, $T_p = 160$ and $T = 1024$

*2) Hard decision-based estimation:* Let $\tilde{\boldsymbol{X}}_d$ be the $n \times T_p$ hard-decision matrix. Let $\theta_{i,k} = Pr\left(x_k = \beta_i | \Lambda_{in}^{det}\right)$ computed from $\boldsymbol{P}_j$. The hard decision on $x_k$ is defined by

$$\tilde{X}_{d,k} = \beta_{i^*} \text{ with } i^* = \arg\max_{1 \le i \le M} \theta_{i,k}. \tag{19}$$

We then propose to update the channel estimation by

$$\hat{\boldsymbol{H}}_i = \left(\boldsymbol{Y}_p \boldsymbol{X}_p^H + \boldsymbol{Y}_d \tilde{\boldsymbol{X}}_d^H\right)\left(\boldsymbol{X}_p \boldsymbol{X}_p^H + \tilde{\boldsymbol{X}}_d \tilde{\boldsymbol{X}}_d^H\right)^{-1}. \tag{20}$$

*3) Soft decision-based estimation:* So as to preserve the information delivered by the FEC decoder, we propose to use $\boldsymbol{\Theta}(t) = (\theta_{i,k}(t))_{1 \le i \le M, 1 \le k \le N}$ the probabilities matrix at time $t$ to compute soft-decisions and update the EM channel estimation as follows:

$$\hat{\boldsymbol{H}}_i = \left(\boldsymbol{Y}_p \boldsymbol{X}_p^H + \sum_{t=T_p+1}^{T} \boldsymbol{y}(t) \boldsymbol{\beta}^H \boldsymbol{\Theta}(t)\right) \tag{21}$$

$$\times \left(\boldsymbol{X}_p \boldsymbol{X}_p^H + \sum_{t=T_p+1}^{T} \boldsymbol{\Theta}^T(t) \boldsymbol{\beta} \boldsymbol{\beta}^H \boldsymbol{\Theta}(t)\right)^{-1},$$

where $\boldsymbol{\beta} = [\beta_1, \ldots, \beta_M]^T$ is the modulation vector.

## V. Simulation Results

We first consider uncoded systems to support the efficiency of the combination of FAS detection with EM estimation update as compared to the MMSE detection-based one. In Fig. 2, we consider $n = N = 64$ and $M = 4$. The EM estimation update is done from hard-decisions taken from detection output. The MSE is plotted after three iterations. The EM estimation combined with FAS detection outperforms the one with MMSE detection with a gain which increases as the SNR gets higher (gain of about $2dB$ at $MSE = 1$ and $3.7dB$ at $MSE = 0.5$).

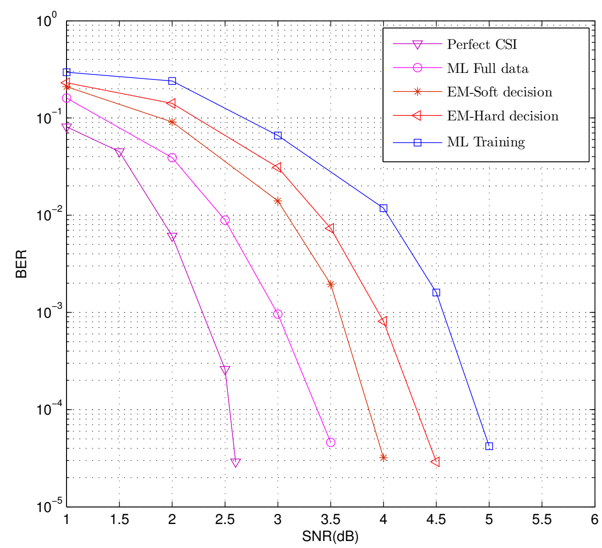


Fig. 4: BER performance with coded 4-QAM, $n = N = 64$, $T_p = 160$ and $T = 1024$.

In the remaining of this section, we consider coded systems with 4-QAM and convolutional code (CC) whose polynomials in octal are $(13, 15)$ (code rate equal to 0.5). A frame consists of 432 short codewords of length equal to 256 coded bits, $T_p = 160$, which makes $T = 1024$.

In Fig. 3, $n = N = 64$. Our purpose is to evaluate the channel estimate accuracy achieved by the turbo receiver (soft decision-based, after six iterations) as compared to the ML

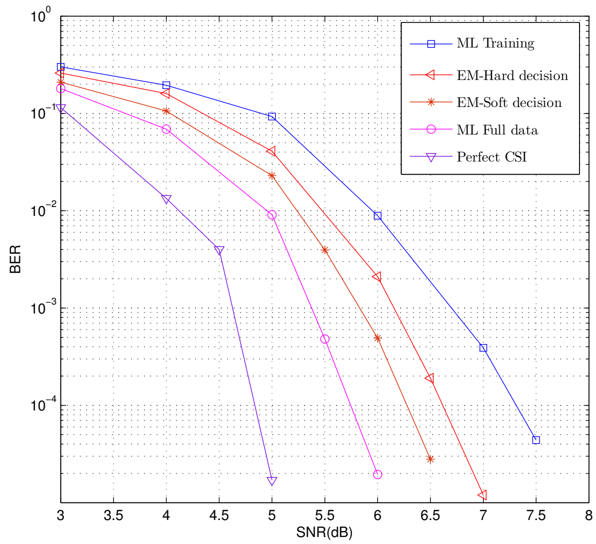


Fig. 5: BER performance with coded 4-QAM, $n = 50$, $N = 64$, $T_p = 160$ and $T = 1024$

training-based estimation and to the lower ML bound referred to as "ML full-data". ML full-data assumes that the whole frame (data and pilot) is known at the receiver and used as training sequence. We observe that the proposed scheme achieves the same MSE as the ML full-data from $SNR = 10dB$. The proposed semi-blind estimation outperforms the ML training-based estimation, with a gain of about $9dB$ at $MSE = 1$. These observations support the efficiency of the proposed estimation.

Then, we compare the two proposed strategies to update the EM channel estimation (hard-decision based and soft-decision based) and we study their impact on the turbo receiver error rate performance. We have also plotted the performance with perfect channel state information, ML full-data estimation and ML training-based estimation.

In Fig. 4, $n = N = 64$. Compared to the perfect CSI knowledge lower bound, at $BER = 10^{-3}$, we observe a loss of 0.75dB for ML-full data, 1.4dB for soft decision-based EM, 1.8dB for hard decision-based EM and 2.5dB for ML-training. The difference between soft and hard decisions-based versions increases slowly with the SNR range and reaches 0.5dB at $BER = 10^{-4}$. In Fig. 5, we consider the underdetermined case with $N = 64$ and $n = 50$. We get roughly the same losses compared to the perfect CSI knowledge as in the determined case. We can thus deduce that the use of detected data to refine the EM channel estimation is efficient as it enables to improve the receiver performance by 0.6-1dB depending on the approach (either hard decision or soft decision-based). Future work will study the best proportion between pilots and data to achieve a compromise between spectral efficiency and transmission quality.

## VI. Conclusion

In this paper we addressed the problem of semi-blind channel estimation in large MIMO FEC-coded systems with finite alphabets assuming limited pilot sequence length. We proposed a turbo FAS-based detection receiver which combines estimation, detection and FEC decoding and we defined two ways of updating the FAS-based EM channel estimation from FEC decoder output. Simulations showed the efficiency of the proposed scheme which performs close to the ML full-data lower bound, with a superiority of the one based on soft decisions. Future work will focus on a compromise between spectral efficiency and target error rate performance as a function of the system dimensions and the ratio between pilots and data.

## References


[1] F. Li, H. Wang, M. Ying, W. Zhang, and J. Lu, “Channel estimations based on superimposed pilots for massive MIMO uplink systems,” in *2016 8th International Conference on Wireless Communications Signal Processing (WCSP)*, Oct 2016, pp. 1–5.

[2] C. Shin, R. W. Heath, and E. J. Powers, “Blind channel estimation for MIMO-OFDM systems,” *IEEE Transactions on Vehicular Technology*, vol. 56, no. 2, pp. 670–685, March 2007.

[3] A. Ladaycia, A. Mokraoui, K. Abed-Meraim, and A. Belouchrani, “What semi-blind channel estimation brings in terms of throughput gain ?” in *2016 10th International Conference on Signal Processing and Communication Systems (ICSPCS)*, Dec 2016, pp. 1–6.

[4] C. H. Aldana, E. de Carvalho, and J. M. Cioffi, “Channel estimation for multicarrier multiple input single output systems using the EM algorithm,” *IEEE Transactions on Signal Processing*, vol. 51, no. 12, pp. 3280–3292, Dec 2003.

[5] M. Abuthinien, S. Chen, and L. Hanzo, “Semi-blind joint maximum likelihood channel estimation and data detection for MIMO systems,” *IEEE Signal Processing Letters*, vol. 15, pp. 202–205, 2008.

[6] R. Otnes and M. Tuchler, “Soft iterative channel estimation for turbo equalization: comparison of channel estimation algorithms,” in *The 8th International Conference on Communication Systems, 2002. ICCS 2002.*, vol. 1, Nov 2002, pp. 72–76 vol.1.

[7] H. Zamiri-Jafarian and S. Pasupathy, “EM-based recursive estimation of channel parameters,” *IEEE Transactions on Communications*, vol. 47, no. 9, pp. 1297–1302, Sep 1999.

[8] E. Nayebi and B. D. Rao, “Semi-blind channel estimation for multiuser massive MIMO systems,” *IEEE Transactions on Signal Processing*, vol. 66, no. 2, pp. 540–553, Jan 2018.

[9] M. Tuchler, R. Otnes, and A. Schmidbauer, “Performance of soft iterative channel estimation in turbo equalization,” in *2002 IEEE International Conference on Communications. Conference Proceedings. ICC 2002 (Cat. No.02CH37333)*, vol. 3, 2002, pp. 1858–1862 vol.3.

[10] X. Wautelet, C. Herzet, A. Dejonghe, J. Louveaux, and L. Vandendorpe, “Comparison of EM-based algorithms for MIMO channel estimation,” *IEEE Transactions on Communications*, vol. 55, no. 1, pp. 216–226, Jan 2007.

[11] Z. Hajji, K. Amis, and A. Aïssa-El-Bey, “Iterative receivers for large-scale MIMO systems with finite-alphabet simplicity-based detection,” *submitted to IEEE Transactions on Communications*, 2018.

[12] ——, “Channel estimation with finite-alphabet simplicity-based detection for large-scale MIMO systems,” *submitted to IEEE Transactions on Communications*, 2018.

[13] Z. Hajji, A. Aïssa-El-Bey, and K. Amis, “Simplicity-based recovery of finite-alphabet signals for large-scale MIMO systems,” *Digital Signal Processing*, vol. 80, pp. 70–82, September 2018.

[14] Y. Nesterov and A. Nemirovskii, *Interior-Point Polynomial Algorithms in Convex Programming*. Society for Industrial and Applied Mathematics, 1994.