Symbol MAP Detection Aided with First Channel Estimate for TDMA Systems Over a Frequency Selective Channel

Inès Kammoun [#1], Mohamed Siala [*2], Noura Sellami [#3], Aline Roumy [†4]

[#] *LETI Laboratory, ENIS, BPW 3038 Sfax- Tunisia*
[1] `ines.kammoun@ieee.org`
[3] `noura.sellami@isecs.rnu.tn`

[*] *Sup'Com, Cité Technologique des communications, Ariana, Tunisia*
[2] `mohamed.siala@supcom.rnu.tn`

[†] *IRISA-INRIA, Campus de Beaulieu, 35042 Rennes Cedex, France*
[4] `Aline.Roumy@irisa.fr`

Abstract—In this paper, we consider joint iterative equalization and channel estimation for TDMA systems over a frequency selective, quasi-static fading channel. We assume a slowly varying channel impulse response from one burst to the following one. We modify the maximum *a posteriori* (MAP) criterion by taking into account an *a priori* channel estimate obtained from the burst preceding the one processed. Assuming that the receiver knows the channel statistics, we propose to use the EM (Expectation-Maximization) algorithm for a joint iterative equalization and channel estimation. During each iteration of the receiver, the EM algorithm reestimates the channel by using the *a posteriori probabilities* (APP) available at the output of the equalizer. The algorithm we propose has linear-time complexity per iteration. Simulations show that our MAP receiver reaches the performance of the MAP equalizer with perfect channel knowledge after a few iterations, for high SNR and for small values of normalized Doppler spread. They show also that our receiver outperforms the MAP equalizer that uses a channel estimate obtained by using a training sequence.

I. INTRODUCTION

In most practical wireless communication systems, the receiver requires an equalizer, whose task is to detect the transmitted symbols in the presence of intersymbol interference (ISI) and channel noise. Most previous works use a suboptimal approach where the equalizer assumes perfect channel knowledge and uses the channel estimate as if it was the true channel. The problem of equalization for non-ideal channel knowledge has been tackled in [3], [4] and [5]. In [3], the sequence MAP criterion for equalization was derived in the case of transmission over a frequency selective channel, for Time Division Multiple Access (TDMA) systems. The derivation of the MAP criterion led in this case to an expression for which no exact implementation exists in the literature. An approximate solution to the problem has been proposed in [4] by using an iterative burst-by-burst equalizer based on the Expectation Maximization (EM) algorithm. In [5], an optimum MAP discrete-time rake receiver has been proposed for CDMA systems when the channel is not perfectly known. In this paper, we propose a new approximate solution

to the problem by performing an optimization on the channel realizations and not on the transmitted sequences. We modify the MAP criterion by taking into account an *a priori* channel estimate obtained from the burst preceding the one processed. We assume a slowly varying channel impulse response from one burst to the following one. Assuming that the receiver knows the channel statistics, we propose to use the EM (Expectation-Maximization) algorithm [6] for a joint iterative equalization and channel estimation. During each iteration of the receiver, the EM algorithm with respect to the MAP criterion reestimates the channel by using the *a posteriori probabilities* (APP) available at the output of the equalizer. The algorithm we propose has linear-time complexity per iteration. The paper is organized as follows. In section 2, we describe the system model. In section 3, we develop the EM algorithm according to the MAP criterion. In section 4, we give simulation results.

II. SYSTEM MODEL

We consider a data transmission system over a frequency selective channel with an impulse response of length L. We assume that transmissions can be organized into blocks of J bursts. Each burst is composed of N data symbols s_n with transmitted energy E_n and time position $p_n = nT, n = 0, 1, ..., N-1$, where T denotes the symbol period. We assume that the first transmitted burst of each block comprises N_p pilot symbols and $N - N_p$ data symbols. These pilot symbols are used for channel and noise statistical characteristics estimation and channel estimation for the first transmitted burst but the following bursts dont require pilot symbols. We assume that the bursts are transmitted over a wide sense stationary uncorrelated scattering (WSSUS) channel. Let B_{Dl} denotes the Doppler spread of the lth path. For simplicity of presentation, we assume that the L paths have the same Doppler spread and denote it by B_D. We assume also that the Doppler spread is sufficiently small that the channel can be assumed to be constant over each burst and to vary from one burst to another.

Let's denote h_l^j the channel coefficient corresponding to the lth path, $l = 0, \cdots, L-1$, of the jth burst, $j = 0, \cdots, J-1$. Within the jth burst, the sample of the received signal at time instant n can be written as

$$r_n^j = \sum_{l=0}^{L-1} h_l^j s_{n-l}^j + w_n^j,\ 0 \le n \le N+L-2 \quad (1)$$

where s_k^j are the transmitted symbols of the jth burst and w_n^j are AWGN complex samples corresponding to the jth burst with probability density function (pdf) $\mathcal{N}(0, \sigma^2)$ and σ^2 is the variance of the noise. The virtual symbols $s_{-1}^j, \ldots, s_{1-L}^j$ are assumed to be known at the receiver and can be set to 0. Let $\mathbf{s}_n^j = \begin{bmatrix} s_n^j & \cdots & s_{n-L+1}^j \end{bmatrix}^T$ and $\mathbf{h}^j = \begin{bmatrix} h_0^j & h_1^j & \cdots & h_{L-1}^j \end{bmatrix}^T$ be the impulse response of the ISI channel.

The received symbol in (1) can be written as

$$r_n^j = \mathbf{s}_n^{j\,T} \mathbf{h}^j + w_n^j,\ 0 \le n \le N+L-2. \quad (2)$$

We introduce, for each transmitted burst, the vector $\mathbf{r}^j = \begin{bmatrix} r_0^j & \cdots & r_{N+L-2}^j \end{bmatrix}^T$ of $N+L-1$ received samples.

III. MODIFIED CRITERION FOR MAXIMUM *a posteriori* CHANNEL ESTIMATION

For a robust intersymbol interference channel estimation and transmitted symbols detection, we propose an EM-based maximum *a posteriori* algorithm. The MAP estimate $\bar{\mathbf{h}}^j$ of the channel $\mathbf{h}^j$ is defined as

$$\bar{\mathbf{h}}^j = \arg\max_{\mathbf{h}} p\left(\mathbf{h} \mid \mathbf{r}^j\right). \quad (3)$$

In this paper, we propose to modify this criterion taking into account the channel estimate $\bar{\mathbf{h}}^{j-1}$ of the burst preceding the one processed as follows

$$\bar{\mathbf{h}}^j = \arg\max_{\mathbf{h}} p\left(\mathbf{h} \mid \mathbf{r}^j, \bar{\mathbf{h}}^{j-1}\right). \quad (4)$$

Our idea is that the first burst of each block of J bursts contains pilot symbols serving for the channel estimation of the first burst and that the estimation of the following bursts is based on an initialization resulting from the estimation of the burst preceding the one processed. Directly solving the equation (4) is an intractable problem. However, the solution can be reached easily by means of the iterative EM (Expectation-Maximization) algorithm. Let's denote $\mathbf{s}^j = \begin{bmatrix} s_{N+L-2}^j & \cdots & s_0^j \end{bmatrix}^T$ the vector of known and unknown symbols transmitted within the $j-th$ burst and characterized by an *a priori* probability distribution $p\left(\mathbf{s}^j\right)$. The EM algorithm inductively reestimates the vector $\mathbf{h}$ so that a monotonic increase in the *a posteriori* conditional pdf $p\left(\mathbf{h} \mid \mathbf{r}, \bar{\mathbf{h}}^{j-1}\right)$ is guaranteed. Given the received vector $\mathbf{r}^j$ and the *a priori* channel estimate $\bar{\mathbf{h}}^{j-1}$, the EM algorithm starts with an initial guess $\mathbf{h}^{j(0)}$ of $\mathbf{h}^j$. The evolution from the estimate $\mathbf{h}^{j(i)}$ to the new estimate $\mathbf{h}^{j(i+1)}$ is performed via the auxiliary function

by carrying the following expectation and maximization steps:
Expectation step: Compute $Q\left(\mathbf{h}, \mathbf{h}^{j(i)}\right)$ given by

$$Q\left(\mathbf{h}, \mathbf{h}^{j(i)}\right) = \sum_{\mathbf{s}^j} p\left(\mathbf{s}^j, \mathbf{h}^{j(i)} \mid \mathbf{r}^j, \bar{\mathbf{h}}^{j-1}\right) \log p\left(\mathbf{s}^j, \mathbf{h} \mid \mathbf{r}^j, \bar{\mathbf{h}}^{j-1}\right) \quad (5)$$

where the latter sum is operated over all possible transmitted data vectors during one block.
Maximization step: Find the re-estimate $\mathbf{h}^{j(i+1)}$ such that

$$\mathbf{h}^{j(i+1)} = \arg\max_{\mathbf{h}} Q\left(\mathbf{h}, \mathbf{h}^{j(i)}\right). \quad (6)$$

We notice that the auxiliary function has several global maxima. This leads to an ambiguity in the MAP estimation of the discrete channel gain factors. To avoid this problem we consider in our study that each block of J bursts comprises pilot symbols which are known by the receiver.
Next, for simplicity of presentation we denote $\mathbf{h}^{(i)}$ instead of $\mathbf{h}^{j(i)}$, $\mathbf{r}$ instead of $\mathbf{r}^j$ and $\mathbf{s}$ instead of $\mathbf{s}^j$. We calculate now the auxiliary function given in (5).

First, by noticing that

$$p\left(\mathbf{s}, \mathbf{h}^{(i)} \mid \mathbf{r}, \bar{\mathbf{h}}^{j-1}\right) = p\left(\mathbf{s} \mid \mathbf{h}^{(i)}, \mathbf{r}, \bar{\mathbf{h}}^{j-1}\right) p\left(\mathbf{h}^{(i)} \mid \mathbf{r}, \bar{\mathbf{h}}^{j-1}\right), \quad (7)$$

$$p\left(\mathbf{s}, \mathbf{h} \mid \mathbf{r}, \bar{\mathbf{h}}^{j-1}\right) = p\left(\mathbf{h} \mid \mathbf{s}, \mathbf{r}, \bar{\mathbf{h}}^{j-1}\right) p\left(\mathbf{s} \mid \mathbf{r}, \bar{\mathbf{h}}^{j-1}\right) \quad (8)$$

and that the terms $p\left(\mathbf{h}^i \mid \mathbf{r}, \bar{\mathbf{h}}^{j-1}\right)$ and $p\left(\mathbf{s} \mid \mathbf{r}, \bar{\mathbf{h}}^{j-1}\right)$ are independent of $\mathbf{h}$, the auxiliary function (5) is equivalent to

$$Q\left(\mathbf{h}, \mathbf{h}^{(i)}\right) = \sum_{\mathbf{s}} p\left(\mathbf{s} \mid \mathbf{h}^{(i)}, \mathbf{r}, \bar{\mathbf{h}}^{j-1}\right) \log\ p\left(\mathbf{h} \mid \mathbf{s}, \mathbf{r}, \bar{\mathbf{h}}^{j-1}\right). \quad (9)$$

By using the conditional independence between $\mathbf{h}$ and $\mathbf{s}$ given $\bar{\mathbf{h}}^{j-1}$, we obtain

$$p\left(\mathbf{h} \mid \mathbf{s}, \mathbf{r}, \bar{\mathbf{h}}^{j-1}\right) = \frac{p\left(\mathbf{r} \mid \mathbf{s}, \mathbf{h}, \bar{\mathbf{h}}^{j-1}\right) p\left(\mathbf{h} \mid \bar{\mathbf{h}}^{j-1}\right)}{p\left(\mathbf{r} \mid \mathbf{s}, \bar{\mathbf{h}}^{j-1}\right)}.$$

Then, the auxiliary function can be written as

$$\begin{aligned}
Q\left(\mathbf{h}, \mathbf{h}^{(i)}\right) &= \sum_{\mathbf{s}} p\left(\mathbf{s} \mid \mathbf{h}^{(i)}, \mathbf{r}, \bar{\mathbf{h}}^{j-1}\right) \log\ p\left(\mathbf{r} \mid \mathbf{s}, \mathbf{h}, \bar{\mathbf{h}}^{j-1}\right) \\
&+ \sum_{\mathbf{s}} p\left(\mathbf{s} \mid \mathbf{h}^{(i)}, \mathbf{r}, \bar{\mathbf{h}}^{j-1}\right) \log p\left(\mathbf{h} \mid \bar{\mathbf{h}}^{j-1}\right) \\
&- \sum_{\mathbf{s}} p\left(\mathbf{s} \mid \mathbf{h}^{(i)}, \mathbf{r}, \bar{\mathbf{h}}^{j-1}\right) \log p\left(\mathbf{r} \mid \mathbf{s}, \bar{\mathbf{h}}^{j-1}\right)
\end{aligned} \quad (10)$$

Since the third term is independent of $\mathbf{h}$, it does not influence the maximization of (6). We notice that given the channel estimate at iteration i we can compute the probability $p\left(\mathbf{s} \mid \mathbf{h}^{(i)}, \mathbf{r}\right)$ independently from $\bar{\mathbf{h}}^{j-1}$. Moreover, from the system model, we can compute $p(\mathbf{r} \mid \mathbf{s}, \mathbf{h})$ independently from $\bar{\mathbf{h}}^{j-1}$. We must then compute and maximize the equivalent form of the auxiliary function:

$$\begin{aligned}
Q\left(\mathbf{h}, \mathbf{h}^{(i)}\right) &= \sum_{\mathbf{s}} p\left(\mathbf{s} \mid \mathbf{h}^{(i)}, \mathbf{r}\right) \log p\left(\mathbf{r} \mid \mathbf{s}, \mathbf{h}\right) \\
&+ \sum_{\mathbf{s}} p\left(\mathbf{s} \mid \mathbf{h}^{(i)}, \mathbf{r}\right) \log p\left(\mathbf{h} \mid \bar{\mathbf{h}}^{j-1}\right)
\end{aligned} \quad (11)$$

which is equivalent to

$$\begin{aligned}
Q\left(\mathbf{h}, \mathbf{h}^{(i)}\right) &= \sum_{\mathbf{s}} p\left(\mathbf{s} \mid \mathbf{h}^{(i)}, \mathbf{r}\right) \log p\left(\mathbf{r} \mid \mathbf{s}, \mathbf{h}\right) \\
&+ \log p\left(\mathbf{h} \mid \bar{\mathbf{h}}^{j-1}\right).
\end{aligned} \quad (12)$$

Next, we assume that the L resolvable paths of the channel are independent. We model the channel coefficient h_l^j corresponding to the lth path, $l = 0, \cdots, L-1$, of the jth burst, $j = 0, \cdots, J-1$ as follows :

$$h_l^j = \alpha_l^j \bar{h}_l^{j-1} + F_l^j, \tag{13}$$

where F_l^j is the innovation and the coefficient α_l^j is to be determined next based on the statistical properties of the innovations process. We model also $\bar{h}_l^{j-1}$ as noisy version of h_l^{j-1}, where the noise distortion has a Gaussian distribution as follows :

$$\bar{h}_l^{j-1} = h_l^{j-1} + b_l^{j-1}. \tag{14}$$

We assume that after estimation of the channel corresponding to the $j-1^{th}$ burst, the sequence of data is correctly detected. That means that the data sequence of the $j-1^{th}$ burst plays completely the role of a reference sequence for the jth burst. Hence, the noise distortion corresponding to the estimation of channel $j-1$ has a Gaussian distribution and we can have a perfect knowledge of its statistics. We notice that this hypothesis is verified for large values of the signal to noise ratio and for large values of N.

Since we assume that the bursts are transmitted over a WSSUS channel, we can easily show that the coefficient α_l^j verifies

$$\alpha_l^j = \frac{R_{hl}(1)}{R_{hl}(0) + R_{bl}^j(0)}, \tag{15}$$

where $R_{hl}(n) = E[(h_l^j)^* h_l^{j+n}]$ and $R_{bl}(0)$ is the variance of b_l, which are independent of j. The operator $(.)^*$ denotes conjugation. We note that α_l^j is independent of j and will be denoted by α_l. We note also that

$$R_{hl}(1) = R_{hl}(0).J_0\left(2\pi B_D T (N+L-1)\right), \tag{16}$$

where $J_0(\cdot)$ is the 0th-order Bessel function of the first kind. We show also that the innovation F_l^j is Gaussian distributed with variance $\theta_l = R_{hl}(0) - 2\alpha_l R_{hl}(1) + \alpha_l^2(R_{hl}(0) + R_{bl})$, which is independent of j.

Then we can write

$$p\left(h_l \mid \bar{h}_l^{j-1}\right) = \frac{1}{\pi\theta_l} \exp\left(-\frac{\left|h_l - \alpha_l \bar{h}_l^{j-1}\right|^2}{\theta_l}\right) \tag{17}$$

Since we assume that the L resolvable paths of the channel are independent, we obtain

$$p\left(\mathbf{h} \mid \bar{\mathbf{h}}^{j-1}\right) = \frac{1}{(\pi^L)\prod_{l=0}^{L-1}\theta_l} \exp\left(-\sum_{l=0}^{L-1}\left(\frac{|h_l - \alpha_l \bar{h}_l^{j-1}|^2}{\theta_l}\right)\right), \tag{18}$$

.

Using the system models (1) and (2), we know that

$$p(\mathbf{r} \mid \mathbf{s}, \mathbf{h}) = \frac{1}{(\pi\sigma^2)^K} \exp\left(-\sum_{k=0}^{K-1}\frac{|r_k - \mathbf{s}_k^T\mathbf{h}|^2}{\sigma^2}\right) \tag{19}$$

where $K = N + L - 1$.

Finally the auxiliary function is given by

$$Q\left(\mathbf{h}, \mathbf{h}^{(i)}\right) = c - \sum_{\mathbf{s}_k} p\left(\mathbf{s}_k \mid \mathbf{h}^{(i)}, \mathbf{r}\right)\left(\sum_{k=0}^{K-1}\frac{|r_k - \mathbf{s}_k^T\mathbf{h}|^2}{\sigma^2}\right) - \left(\mathbf{h} - \mathbf{\Lambda}\bar{\mathbf{h}}_l^{j-1}\right)^\dagger \mathbf{\Xi}^{-1}\left(\mathbf{h} - \mathbf{\Lambda}\bar{\mathbf{h}}_l^{j-1}\right), \tag{20}$$

where c is an additive factor independent of $\mathbf{h}$, $\mathbf{\Lambda} = \text{diag}\{\alpha_l\}$ and $\mathbf{\Xi} = \text{diag}\{\theta_l\}$. Maximizing $Q(\mathbf{h}, \mathbf{h}^i)$ as a function of $\mathbf{h}$ yields

$$\mathbf{h}^{(i+1)} = \left(\sum_{k=L-1}^{K-1} r_k E\left(\mathbf{s_k} \mid \mathbf{h}^{(i)}, \mathbf{r}\right) + \sigma^2 \mathbf{\Xi}^{-1}\mathbf{\Lambda}\bar{\mathbf{h}}_l^{j-1}\right) \left(\sum_{k=L-1}^{K-1} E\left(\mathbf{s}_k\mathbf{s}_k^\dagger \mid \mathbf{h}^{(i)}, \mathbf{r}\right) + \sigma^2 \mathbf{\Xi}^{-1}\right)^{-1}, \tag{21}$$

where

$$E\left(\mathbf{s}_k \mid \mathbf{h}^{(i)}, \mathbf{r}\right) = \sum_{\mathbf{d}\in\mathbf{D}} \mathbf{d}\, p\left(\mathbf{s}_k = \mathbf{d} \mid \mathbf{h}^{(i)}, \mathbf{r}\right) \tag{22}$$

and

$$E\left(\mathbf{s}_k\mathbf{s}_k^\dagger \mid \mathbf{h}^{(i)}, \mathbf{r}\right) = \sum_{\mathbf{d}\in\mathbf{D}} \mathbf{d}\mathbf{d}^\dagger\, p\left(\mathbf{s}_k = \mathbf{d} \mid \mathbf{h}^{(i)}, \mathbf{r}\right). \tag{23}$$

Here, $\mathbf{D}$ denotes the set of possible vectors $\mathbf{s}_k$.

The conditional probabilities $p\left(\mathbf{s}_k = \mathbf{d} \mid \mathbf{h}^{(i)}, \mathbf{r}\right)$ can be computed by means of the forward-backward algorithm [2]. This algorithm introduces the functions:

$$\alpha_k(m) = p\left(\Sigma_k = m, \{r_n\}_{n=0}^{k-1} \mid \mathbf{h}^{(i)}\right), \tag{24}$$

and

$$\beta_k(m) = p\left(\{r_n\}_{n=k}^{K-1} \mid \Sigma_k = m, \mathbf{h}^{(i)}\right), \tag{25}$$

Σ_k denoting the state of the trellis at time k and $m = 0, ..., M-1$ the M distinct states of the code trellis. These functions can be computed respectively by the forward recursion

$$\alpha_{k+1}(m) = \sum_{m'=0}^{M-1} \alpha_k(m')\gamma_k(m', m), \tag{26}$$

with the boundary conditions

$$\alpha_0(0) = 1, \alpha_0(m) = 0, \text{ for } m \neq 0 \tag{27}$$

and the backward recursion

$$\beta_k(m) = \sum_{m'=0}^{M-1} \beta_{k+1}(m')\gamma_{k+1}(m', m), \tag{28}$$

with the boundary conditions

$$\beta_{K-1}(0) = 1, \beta_{K-1}(m) = 0 \text{ for } m \neq 0. \tag{29}$$

The conditional probability function

$$\gamma_k(m', m) = p\left(\Sigma_k = m, r_k \mid \Sigma_{k-1} = m', \mathbf{h}^{(i)}\right) \tag{30}$$

can be computed as follows

$$\gamma_k(m', m) = \sum_{\mathbf{d}} p_k\left(m \mid m'\right) q_k\left(\mathbf{d} \mid m', m\right) p\left(\mathbf{r} \mid \mathbf{s}_k = \mathbf{d}, \mathbf{h}^{(i)}\right) \tag{31}$$

where $p_k(m \mid m') = p(\Sigma_k = m \mid \Sigma_{k-1} = m')$ and $q_k(\mathbf{d} \mid m', m) = p(\mathbf{s}_k = \mathbf{d} \mid \Sigma_{k-1} = m', \Sigma_k = m)$. The needed probabilities are given by

$$p(\mathbf{s}_k = \mathbf{d} \mid \mathbf{h}^{(i)}, \mathbf{r})$$
$$= \sum_{m,m'} \alpha_{k-1}(m')\gamma_k(m', m)\beta_k(m)q_k(\mathbf{d} \mid m', m) \quad (32)$$

IV. IMPLEMENTATION AND COMPLEXITY EVALUATION

We give here a summary of the algorithm and evaluate the complexity of its different steps :

1) Initialization. The channel estimate of the first transmitted burst of each block of J bursts is obtained using the classical MAP criterion (3) which gives

$$\mathbf{h}^{0(i+1)} = \left(\sum_{k=L-1}^{K-1} r_k^0 E\left(\mathbf{s}_k \mid \mathbf{h}^{0(i)}, \mathbf{r}^0\right) \right)$$
$$\left(\sum_{k=L-1}^{K-1} E\left(\mathbf{s}_k \mathbf{s}_k^\dagger \mid \mathbf{h}^{0(i)}, \mathbf{r}^0\right) \right)^{-1}, \quad (33)$$

where $E\left(\mathbf{s}_k \mid \mathbf{h}^{0(i)}, \mathbf{r}^0\right)$ and $E\left(\mathbf{s}_k(\mathbf{s}_k)^\dagger \mid \mathbf{h}^{0(i)}, \mathbf{r}^0\right)$ are given respectively in (22) and (23), by replacing $\mathbf{h}^{(i)}$ by $\mathbf{h}^{0(i)}$ and $\mathbf{r}$ by $\mathbf{r}^0$.

We propose to obtain $\mathbf{h}^{0(0)}$ by using a training sequence of length $N_p + L - 1$. Since the channel noise is additive, white and Gaussian, the maximum likelihood estimator based on the observations generated by the training sequence is the least squares estimator [8]. Let $\widetilde{\mathbf{s}} = \begin{bmatrix} s_{n_0+N_p+L-2} & \cdots & s_{n_0} \end{bmatrix}^T$ be the vector of training symbols, with $0 \le n_0 \le N - N_p - L + 1$ and $\widetilde{\mathbf{S}}$ be the $N_p \times L$ Hankel matrix with first column $\begin{bmatrix} s_{n_0+N_p+L-2} & \cdots & s_{n_0+L-1} \end{bmatrix}^T$ and last row $\begin{bmatrix} s_{n_0+L-1} & \cdots & s_{n_0} \end{bmatrix}^T$. The least squares channel estimate $\mathbf{h}^{0(0)}$ is given by

$$\mathbf{h}^{0(0)} = (\widetilde{\mathbf{S}}^\dagger \widetilde{\mathbf{S}})^{-1} \widetilde{\mathbf{S}}^\dagger \widetilde{\mathbf{r}}. \quad (34)$$

The receiver needs to estimate $R_{bl}(0)$. We can easily justify that $R_{bl}(0)$ can be estimated with σ^2/N_p. In fact, the channel estimate $\mathbf{h}^{0(0)}$ can be modelled as $\mathbf{h}^{0(0)} = \mathbf{h}^0 + \mathbf{b}^0$, with $\mathbf{b}^0 = \left(\widetilde{\mathbf{S}}^\dagger \widetilde{\mathbf{S}}\right)^{-1} \widetilde{\mathbf{S}}^\dagger \widetilde{\mathbf{w}}$ being circularly symmetric complex Gaussian distributed with covariance matrix $\mathbf{R}_b^0 = \sigma^2 \left(\widetilde{\mathbf{S}}^\dagger \widetilde{\mathbf{S}}\right)^{-1}$. Since we use a CAZAC sequence, we have $\mathbf{R}_b^0 = \sigma^2/N_p \mathbf{I}_L$. By considering that the SNR is high we can assume that there are no errors of decisions. Then, each burst can be considered as a training sequence for the following burst and equation (34) can be used by replacing the N_p-length training vector by the total sequence of the burst preceding the one processed. Therefore, for a large number N, we can approximate $R_{bl}(j)$ by σ^2/N, for $j \ge 1$. Hence, we have $\mathbf{R}_b^j = \mathbf{R}_b^0 N_0/N$ for $j \ge 1$.

The estimate of the channel noise variance $\widehat{\sigma}^2$ is obtained using the $N_0 + L - 1$ training symbols

$$\widehat{\sigma}^2 = \frac{1}{N_p} \sum_{k=n_0+L-1}^{n_0+N_p+L-2} \left| r_k - \mathbf{s}_k^T \bar{\mathbf{h}}^{(0)} \right|^2. \quad (35)$$

Each decoded sequence in each burst can also serve for having better estimation of σ^2, but the algorithm is not very sensitive to the errors of estimation of σ^2. Hence, we estimate σ^2 using the training sequence included in the first burst of every block.

2) For each iteration index i

   a) Using the *a posteriori* probabilities (APP) $p(\mathbf{s}_k = \mathbf{d} \mid \mathbf{h}^{(i)}, \mathbf{r})$ computed by means of the forward-backward algorithm, calculate $E(\mathbf{s}_k \mid \mathbf{h}^{(i)}, \mathbf{r})$ and $E(\mathbf{s}_k(\mathbf{s}_k)^\dagger \mid \mathbf{h}^{(i)}, \mathbf{r})$ defined in (22) and (23), respectively.

   b) Compute a new channel estimation $\mathbf{h}^{(i+1)}$ as defined in (21).

Notice that the inversions and multiplications of the matrices in the initialization step have complexities which do not depend on N. Indeed, the inversions of matrices $\mathbf{R}_h$ and $\left(\sum_{k=L-1}^{K-1} E\left(\mathbf{s}_k(\mathbf{s}_k)^\dagger \mid \mathbf{h}^{(i)}, \mathbf{r}\right) + \sigma^2 \mathbf{R}_h^{-1} \right)$ in equation (21) requires $L^3/6$ multiplications each. Notice that when K is sufficiently high, we can approximate this matrix by a diagonal one which is independent from the transmitted sequence. Hence, for complexity reduction, we propose to compute this matrix only for the first burst of each block. We notice that L is small enough with respect to N and that when N increases the calculation of the covariance matrix remains linear in N but the matrix inversion has always a bounded complexity depending only on the number of paths L.

V. SIMULATION RESULTS

Here we present some simulation results in the case of channel length $L = 3$. We consider block-by-block channel estimation using BPSK symbols. Each burst of $N = 512$ information bits is transmitted over a quasi stationary channel. The different channel taps are modelled as independent zero mean complex Gaussian random variables with variance $1/L$. In Figure 1, we show the performance of the proposed receiver after four iterations and a normalized Doppler spread value $B_D T = 1/100$. In our simulations we assume that the first transmitted burst uses $N_p = 16$ pilot symbols. In order to avoid the propagation of estimation errors we assume that after each J transmitted bursts a burst is transmitted with pilot symbols. We consider the three values $J = 40$, 20 and 10. The signal to noise ratio is given by $SNR = E_b/\sigma^2$ where E_b is the energy per transmitted bit. Comparison is done with the MAP equalizer that has perfect channel knowledge and the MAP equalizer that uses a channel estimate obtained by using 16 pilot symbols included in each transmitted burst.

This figure shows that the best achievable BER for the three values of J considered is almost reached after 4 iterations. We also that the performance of the proposed algorithm, for the three values of J considered, approaches the performance obtained when the equalizer has perfect channel knowledge for large values of SNR. However, we notice a significant degradation in performance for low SNR and for large number J of bursts within one block. For a further characterization of the proposed receiver we show in Figure 2 the BER per-

formance for different values of normalized Doppler spreads when $J = 10$. This figure shows a degradation in performance for large normalized Doppler spreads and for low SNR. For an additional characterization of our receiver, we show in Figure 3 the BER as a function of the SNR for two different numbers of pilot symbols $N_p = 16$ and $N_p = 8$ when $B_D T = 1/100$. Two numbers of bursts per block are considered $J = 20$ and $J = 10$. Simulations also show that when $N_p = 16$, the performance of the proposed receiver at the third iteration approaches the performance obtained when the equalizer has perfect channel knowledge. However, when $N_p = 8$, there is a significant gap in terms of performance compared to the case of perfect channel knowledge.
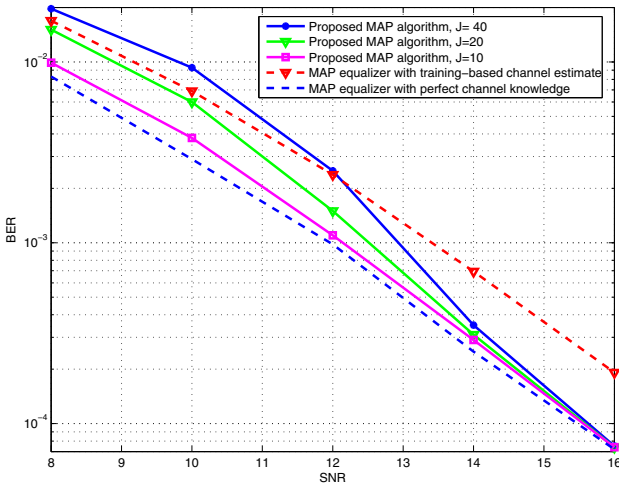

Fig. 3.   BER performance for different numbers of pilot symbols

Fig. 1.   Raw BER as a function of SNR when $B_D T = 1/100$

criterion taking into account the channel estimate of the burst preceding the one processed. We assumed that the channel is invariant during the transmission of one burst and that it vary from one burst to the following one depending on the Doppler spread. In order to avoid the propagation of estimation errors, we assumed that a burst comprising a training sequence is transmitted once every block of J bursts. Simulation results have shown that our iterative receiver can reach after a few iterations the performance of the MAP equalizer with perfect channel knowledge for high SNR. For low SNR, a degradation in performance is noticed when the number J or the normalized Doppler spread value is large. Hence, in practice, it will be interesting to adapt the number of bursts within one block depending on the SNR value or on the doppler spread.
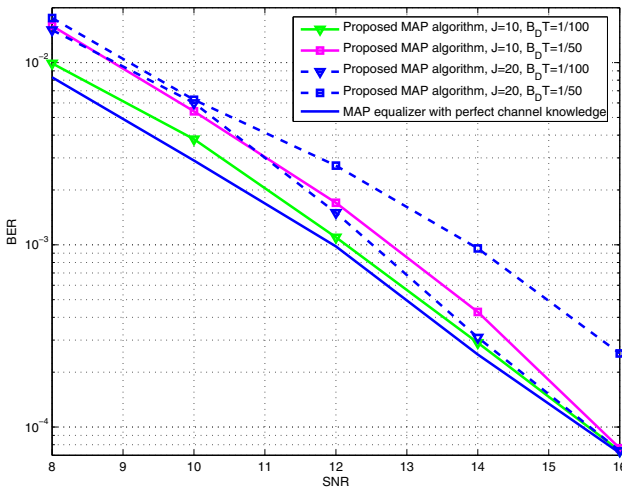

Fig. 2.   Raw BER as a function of SNR for different normalized Doppler spread values

VI. Conclusion

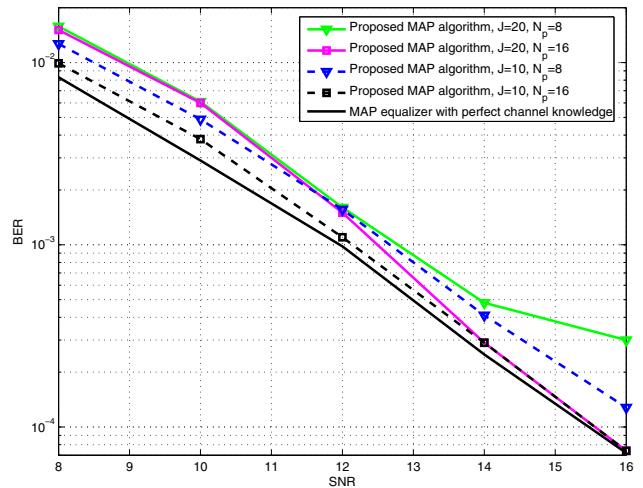In this paper, we proposed an iterative channel estimation algorithm according to a modified maximum *a posteriori*

References

[1] G. D. Forney, "Maximum-likelihood sequence estimation for digital sequences in the presence of intersymbol interference," *IEEE Trans. Inf. Theory, vol. 18, pp. 363-378*, May 1972.
[2] L. R. Bahl, J. Cocke, F. Jelinek, and J. Raviv, "Optimal decoding of linear codes for minimizing symbol error rate," *IEEE Trans. Inf. Theory, vol. IT-32, pp.284-287*, March 1974.
[3] M. T. Tuchler and M. Mecking, "Equalization for non-ideal channel knowledge," *Conf. on Inf. Sciences and Systems*, March 2003.
[4] N. Sellami, M. Siala, A. Roumy and I. Kammoun, "Generalized MAP: sequence detection for non-ideal frequency selective channel knowledge," *IEEE International Conference on Acoustics, Speech and Signal Processing, ICASSP 2007*, Honolulu USA, April 2007.
[5] M. Siala, "Maximum *a posteriori* Decorrelating Discrete-Time Rake Receiver," *Annals of Telecommunications, vol. 59, n. 3-4*, March-April 2004.
[6] A. P. Dempster, N. M. Laird, and D. B. Rubin, "Maximum likelihood from incomplete data via the EM algorithm," *Journ. of the Royal Stat. Soc.*, vol. 39, pp. 1-38, 1977.
[7] A. Milewsky, "Periodic sequences with optimal properties for channel estimation and fast startup equalization," *IBM Journal of Research and Development*, vol. 27, no.5, September 1983.
[8] S. Crozier, D. Falconer, and S. Mahmoud, "Least sum of squared errors (LSSE) Mchannel estimation," *IEE Proceedings, vol. 138, pp.371-378*, August 1991.