MAP sequence equalization for imperfect frequency selective channel knowledge *

Noura Sellami[1], Mohamed Siala[2], Aline Roumy[3], and Ines Kammoun[1]

[1] ENIS, Route de Soukra km 4, 3038 Sfax, Tunisia

Email: noura.sellami@isecs.rnu.tn, ines.kammoun@ieee.org

Tel: (216)74274088, Fax: (216)74275595

[2] SUP'COM, Cité Technologique des Communications,

2083 El Ghazala Ariana, Tunisia

Email: mohamed.siala@supcom.rnu.tn, Tel: (216)71857000, Fax: (216)71856829

[3] IRISA-INRIA, Campus de Beaulieu, 35042 Rennes Cedex, France

Email: aline.roumy@irisa.fr, Tel: (33)299847394, Fax: (33)299847171

Corresponding author: Noura Sellami

Abstract

This paper considers the problem of Maximum *a Posteriori* (MAP) sequence equalization over a frequency selective channel when the channel is not perfectly known at the receiver. The imperfect channel estimate is provided by a least squares estimator using a training sequence (TS). The derivation of the MAP criterion taking into account the channel estimation error exists in the literature but no direct practical implementation has been proposed. In this paper, we propose an approximate solution to the problem by using an iterative burst-by-burst equalizer based on the Expectation Maximization (EM) algorithm. Simulation results show that the performance of our iterative equalizer is significantly better than the performance of the non iterative equalizer which does not take into account the channel estimation error.

Keywords: MAP equalization, channel estimation, EM algorithm, frequency selective channel, turbo-detection, Soft Output Viterbi Algorithm.

1 Introduction

To combat the effects of intersymbol interference (ISI), an equalizer has to be used. In practice, the receiver does not know perfectly the channel and has to estimate it. In this paper, we consider the case where the equalizer has an imperfect channel estimate provided by the channel estimator which uses a training sequence (TS). 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 [1, 2]. In [1], Tüchler *et al.* derived the Maximum *A Posteriori* (MAP) criterion in the case of transmission over a frequency selective channel, for Time Division Multiple Access (TDMA) systems. They suggested to use sequential decoding or sphere decoding to solve it, since it was not feasible for a direct practical implementation. In [2], an optimum MAP discrete-time rake receiver has been proposed for Code Division Multiple Access (CDMA) systems when the channel is not perfectly known. In this case, the algorithm can be implemented since the Probability Density Function (PDF) involved in the derivation of the MAP criterion can be factored unlike the PDF in [1].

In this paper, we consider, as in [1], a TDMA data transmission system over a frequency selective channel. We propose an approximate solution to the problem of MAP equalization with imperfect channel knowledge using an iterative burst-by-burst equalizer based on the Expectation Maximization (EM) algorithm [3]. The EM algorithm is a general procedure for iterative maximum likelihood or MAP estimation. When the channel and the transmitted symbols are unknown, two applications of the EM algorithm can be considered. Indeed, it can be used to iteratively estimate the channel [4, 5] or to iteratively estimate the transmitted symbols [6, 7, 8]. Here, we propose a new approach using the EM algorithm to give an approximate solution to the intractable problem introduced in [1]. The EM algorithm alternates between two steps, an Expectation (E) step and a Maximization (M) step. In our case, the Maximization step of the EM algorithm is performed by a Viterbi algorithm. In order to include our equalizer into a turbo-detector, we propose a soft output version of it, based on the Soft Output Viterbi Algorithm (SOVA) [9]. Simulation results show that our iterative equalizer performance is significantly better than the performance obtained by using the approximate solution proposed in [1] and than that of the suboptimal non iterative equalizer which does not take into account the channel estimation error.

Throughout this paper, scalars are lower case, and vectors and matrices are bold lower and upper case, respectively. The operators $(.)^T$, $(.)^*$ and $(.)^\dagger$ denote respectively transposition, conjugation and transconjugation. The $L \times L$ identity matrix is denoted by $\mathbf{I}_L$.

2 System model

The input information bit sequence is mapped to the symbol alphabet $\mathcal{A}$. We assume that the transmission is organized into bursts of K symbols each. The transmitted symbols are equiprobable. The channel is supposed to be invariant during one burst and can change from burst to burst. Within one burst, the baseband signal sampled at the symbol rate and received at time k is

$$r_k = \sum_{l=0}^{L} h_l s_{k-l} + w_k = \mathbf{s}_k^T \mathbf{h} + w_k, \quad \forall\, 0 \le k \le K + L - 1$$

where L is the channel memory length, s_k, for $0 \le k \le K - 1$, are the transmitted symbols with average energy E_S and $\mathbf{s}_k = [s_k, \cdots, s_{k-L}]^T$. The virtual symbols s_k, for $-L \le k \le -1$ and $K \le k \le K + L - 1$, are assumed to be known at the receiver and

can be assumed to be 0. In this expression, w_k, for $0 \leq k \leq K + L - 1$, are Additive White Gaussian Noise (AWGN) complex samples with PDF $\mathcal{N}_C(0, \sigma^2)$, where $\sigma^2 = N_0$ and $\mathcal{N}_C(\alpha, \sigma^2)$ denotes a circular complex Gaussian distribution with mean α and variance σ^2. The channel realization vector $\mathbf{h} = [h_0 \ldots h_L]^T$ is circularly symmetric complex Gaussian distributed with PDF $\mathcal{N}_C(\mathbf{0}, \mathbf{R}_h)$, where $\mathbf{R}_h$ is the channel covariance matrix. The PDF of $\mathbf{h}$ is given by

$$p(\mathbf{h}) = \frac{1}{\det(\pi \mathbf{R}_h)} \exp\left(-\mathbf{h}^\dagger \mathbf{R}_h^{-1} \mathbf{h}\right). \tag{1}$$

Let $\mathbf{s} = [s_{K-1}, ..., s_{-L}]^T$ be the $(L + K)$-long vector of symbols and $\mathbf{S}$ be the associated $(K + L) \times (L + 1)$ Hankel matrix with last column $[s_{K-1}, \cdots, s_{-L}]^T$ and last row $[s_0, \cdots, s_{-L}]$. The received signal model can be rewritten in a matrix form as $\mathbf{r} = \mathbf{Sh} + \mathbf{w}$, with $\mathbf{r} = [r_{K+L-1}, ..., r_0]^T$ and $\mathbf{w} = [w_{K+L-1}, ..., w_0]^T$. It follows that:

$$
\begin{aligned}
p(\mathbf{r}|\mathbf{h}, \mathbf{s}) &= \frac{1}{(\pi\sigma^2)^{K+L}} \exp\left(-\sum_{k=0}^{K+L-1} \frac{|r_k - \mathbf{s}_k^T \mathbf{h}|^2}{\sigma^2}\right) \\
&= \frac{1}{(\pi\sigma^2)^{K+L}} \exp\left(-\frac{(\mathbf{r} - \mathbf{Sh})^\dagger (\mathbf{r} - \mathbf{Sh})}{\sigma^2}\right). \tag{2}
\end{aligned}
$$

The initial channel estimate $\hat{\mathbf{h}}$ is provided to the equalizer by a least squares estimator using a training sequence (TS) $\tilde{\mathbf{s}} = [s_{t_0+T_0+L-1}, ..., s_{t_0}]^T$ of length $T_0 + L$, with $0 \leq t_0 \leq K - T_0 - L$ [10]. We assume that the training symbols have the same average energy E_S as the transmitted information data symbols. Let $\tilde{\mathbf{S}}$ be the $T_0 \times (L + 1)$ Hankel matrix with first column $[s_{t_0+T_0+L-1}, \cdots, s_{t_0+L}]^T$ and last row $[s_{t_0+L}, \cdots, s_{t_0}]$. The channel estimate $\hat{\mathbf{h}}$ can then be modeled as $\hat{\mathbf{h}} = \mathbf{h} + \mathbf{b}$, with $\mathbf{b}$ being circularly symmetric complex Gaussian distributed with PDF $\mathcal{N}_C(\mathbf{0}, \mathbf{R}_b)$ and $\mathbf{R}_b = \sigma^2 \left(\tilde{\mathbf{S}}^\dagger \tilde{\mathbf{S}}\right)^{-1}$. It follows that

$$p(\hat{\mathbf{h}}|\mathbf{h}) = \frac{1}{\det(\pi \mathbf{R}_b)} \exp\left(-(\hat{\mathbf{h}} - \mathbf{h})^\dagger \mathbf{R}_b^{-1} (\hat{\mathbf{h}} - \mathbf{h})\right) \tag{3}$$

and

$$p(\hat{\mathbf{h}}) = \frac{1}{\det(\pi (\mathbf{R}_h + \mathbf{R}_b))} \exp\left(-\hat{\mathbf{h}}^\dagger (\mathbf{R}_h + \mathbf{R}_b)^{-1} \hat{\mathbf{h}}\right). \tag{4}$$

3 Problem statement

When the channel is perfectly known at the receiver, the data estimate according to the MAP sequence criterion is given by

$$\hat{\mathbf{s}} = \arg \max_{\mathbf{s}} P(\mathbf{s}|\mathbf{r}, \mathbf{h}), \tag{5}$$

which can be solved efficiently by the Viterbi algorithm [11]. Here, however, an estimate $\hat{\mathbf{h}}$ but not the true channel $\mathbf{h}$ is available at the receiver. Replacing $\mathbf{h}$ by $\hat{\mathbf{h}}$ in the equalization algorithm leads to a loss in signal to noise ratio (SNR) for which a closed form expression is derived in [12]. This loss is significant when the channel estimate is noisy (the TS is short). Hence, in this paper, we consider the exact MAP criterion that takes into account the channel estimation error:

$$\hat{\mathbf{s}} = \arg \max_{\mathbf{s}} P(\mathbf{s}|\mathbf{r}, \hat{\mathbf{h}}). \tag{6}$$

This criterion was first derived in [1] but no efficient implementation was proposed yet.

4 Proposed approximate solution: EM algorithm

We propose to use the EM algorithm [3, 13] to iteratively approach the solution of the problem (6) with a reasonable complexity. Each iteration of the algorithm consists in two steps:

$$E \quad : \quad Q\left(\mathbf{s}, \mathbf{s}^i\right) = \int \log P\left(\mathbf{s}|\mathbf{r}, \mathbf{h}, \hat{\mathbf{h}}\right) p\left(\mathbf{h}|\mathbf{r}, \hat{\mathbf{h}}, \mathbf{s}^i\right) d\mathbf{h} \tag{7}$$

$$M \quad : \quad \mathbf{s}^{i+1} = \arg\max_{\mathbf{s}} Q\left(\mathbf{s}, \mathbf{s}^i\right) \tag{8}$$

where E stands for the Expectation step and M for the Maximization step of the EM algorithm. When the EM algorithm is used to estimate continuous parameters, it was proved in [13] that it converges to a local maximum. However, for discrete parameter estimation, as it is the case in our work, the EM algorithm is not guaranteed to converge to a maximum [14]. Thus, the solution we propose is an approximate solution to the problem introduced in [1]. In the following, we detail a possible implementation of the EM algorithm.

4.1 Expectation step

By noticing that $P\left(\mathbf{s}|\mathbf{r}, \mathbf{h}, \hat{\mathbf{h}}\right) = p\left(\mathbf{r}|\mathbf{s}, \mathbf{h}\right) P\left(\mathbf{s}\right) / p\left(\mathbf{r}|\mathbf{h}\right)$, the function Q in (7), at the $(i+1)^{th}$ iteration, can be simplified as:

$$Q\left(\mathbf{s}, \mathbf{s}^i\right) = \int \log p\left(\mathbf{r}|\mathbf{s}, \mathbf{h}\right) p\left(\mathbf{h}|\mathbf{r}, \hat{\mathbf{h}}, \mathbf{s}^i\right) d\mathbf{h} + \log P\left(\mathbf{s}\right) \underbrace{- \int \log p\left(\mathbf{r}|\mathbf{h}\right) p\left(\mathbf{h}|\mathbf{r}, \hat{\mathbf{h}}, \mathbf{s}^i\right) d\mathbf{h}}_{\alpha}.$$

Since α is independent of $\mathbf{s}$, we focus on the computation of $\int \log p\left(\mathbf{r}|\mathbf{s}, \mathbf{h}\right) p\left(\mathbf{h}|\mathbf{r}, \hat{\mathbf{h}}, \mathbf{s}^i\right) d\mathbf{h}$. Using the independence between $(\mathbf{h}, \hat{\mathbf{h}})$ and $\mathbf{s}^i$, and between $\mathbf{r}$ and $\hat{\mathbf{h}}$ given $(\mathbf{h}, \mathbf{s}^i)$, the Bayes' rule reads:

$$p\left(\mathbf{h}|\mathbf{r}, \hat{\mathbf{h}}, \mathbf{s}^i\right) = \frac{p(\mathbf{r}|\mathbf{h}, \mathbf{s}^i)p(\mathbf{h}|\hat{\mathbf{h}})}{p(\mathbf{r}|\hat{\mathbf{h}}, \mathbf{s}^i)}. \tag{9}$$

We proceed by computing $p(\mathbf{h}|\hat{\mathbf{h}})$. Again, using the Bayes' rule and the expressions of the PDFs (1), (3) and (4), we obtain

$$p(\mathbf{h}|\hat{\mathbf{h}}) \quad = \quad \frac{p(\hat{\mathbf{h}}|\mathbf{h})p(\mathbf{h})}{p(\hat{\mathbf{h}})}$$

$$= \quad \frac{1}{\det(\pi\boldsymbol{\Sigma})} \exp\left(-(\mathbf{h} - \mathbf{u})^\dagger \boldsymbol{\Sigma}^{-1}(\mathbf{h} - \mathbf{u})\right), \tag{10}$$

with $\boldsymbol{\Sigma}^{-1} = \mathbf{R}_b^{-1} + \mathbf{R}_h^{-1}$ and $\mathbf{u} = \boldsymbol{\Sigma}\mathbf{R}_b^{-1}\hat{\mathbf{h}}$. The derivation of $p(\mathbf{r}|\hat{\mathbf{h}}, \mathbf{s}^i)$ needs an additional expectation since

$$p(\mathbf{r}|\hat{\mathbf{h}}, \mathbf{s}^i) = \int p(\mathbf{r}|\mathbf{h}, \mathbf{s}^i)p(\mathbf{h}|\hat{\mathbf{h}})d\mathbf{h}.$$

Using (2) and (10), we obtain

$$p(\mathbf{r}|\hat{\mathbf{h}}, \mathbf{s}^i) = \frac{1}{\det(\pi\mathbf{R}_c^i)} \exp\left(-(\mathbf{r} - \mathbf{c}^i)^\dagger \mathbf{R}_c^{i^{-1}}(\mathbf{r} - \mathbf{c}^i)\right), \tag{11}$$

with

$$\mathbf{R}_c^i = \sigma^2 \mathbf{I}_{K+L} + \mathbf{S}^i \mathbf{\Sigma} \mathbf{S}^{i\dagger} \qquad (12)$$

$$\mathbf{c}^i = \mathbf{R}_c^i \frac{\mathbf{S}^i}{\sigma^2} \left(\frac{1}{\sigma^2} \mathbf{S}^{i\dagger} \mathbf{S}^i + \mathbf{\Sigma}^{-1} \right)^{-1} \mathbf{\Sigma}^{-1} \mathbf{u}, \qquad (13)$$

$\mathbf{S}^i$ being the $(K+L) \times (L+1)$ Hankel matrix associated to $\mathbf{s}^i$ defined as $\mathbf{S}$. Finally, by using the Bayes' rule in (9) and the PDFs in (2), (10) and (11), the desired PDF is given by:

$$p\left(\mathbf{h}|\mathbf{r}, \hat{\mathbf{h}}, \mathbf{s}^i\right) = \frac{1}{\det(\pi \mathbf{R}_d^i)} \exp\left(-(\mathbf{h} - \mathbf{d}^i)^\dagger \mathbf{R}_d^{i^{-1}} (\mathbf{h} - \mathbf{d}^i)\right) \qquad (14)$$

with

$$\mathbf{R}_d^{i^{-1}} = \frac{1}{\sigma^2} \mathbf{S}^{i\dagger} \mathbf{S}^i + \mathbf{R}_b^{-1} + \mathbf{R}_h^{-1} \qquad (15)$$

$$\mathbf{d}^i = \mathbf{R}_d^i \left(\frac{1}{\sigma^2} \mathbf{S}^{i\dagger} \mathbf{r} + \mathbf{R}_b^{-1} \hat{\mathbf{h}} \right). \qquad (16)$$

We have

$$\int \log p\left(\mathbf{r}|\mathbf{s}, \mathbf{h}\right) p\left(\mathbf{h}|\mathbf{r}, \hat{\mathbf{h}}, \mathbf{s}^i\right) d\mathbf{h} = -\frac{1}{\sigma^2} \sum_{k=0}^{K+L-1} \int |r_k - \mathbf{s}_k^T \mathbf{h}|^2 p\left(\mathbf{h}|\mathbf{r}, \hat{\mathbf{h}}, \mathbf{s}^i\right) d\mathbf{h} - (K+L) \log(\pi \sigma^2).$$

The terms of the sum can be computed using (14) and we obtain

$$\int |r_k - \mathbf{s}_k^T \mathbf{h}|^2 p\left(\mathbf{h}|\mathbf{r}, \hat{\mathbf{h}}, \mathbf{s}^i\right) d\mathbf{h} = |r_k - \mathbf{s}_k^T \mathbf{d}^i|^2 + \mathbf{s}_k^T \mathbf{R}_d^i \mathbf{s}_k^*.$$

Finally, the function Q is given by:

$$Q\left(\mathbf{s}, \mathbf{s}^i\right) = -\frac{1}{\sigma^2} \sum_{k=0}^{K+L-1} \left(|r_k - \mathbf{s}_k^T \mathbf{d}^i|^2 + \mathbf{s}_k^T \mathbf{R}_d^i \mathbf{s}_k^* \right) - (K+L) \log(\pi \sigma^2) + \log P(\mathbf{s}) - \alpha. \quad (17)$$

Remark 1: Notice that the channel estimation $\mathbf{d}^i$ obtained while deriving the EM algorithm is the mean of the *a posteriori* probability $p\left(\mathbf{h}|\mathbf{r}, \hat{\mathbf{h}}, \mathbf{s}^i\right)$. This estimation is thus the Minimum Mean Square Estimation of the channel.

4.2   Maximization step

From (17), the maximization step (8) can be performed recursively by using the Viterbi algorithm with branch metric $\left(|r_k - \mathbf{s}_k^T \mathbf{d}^i|^2 + \mathbf{s}_k^T \mathbf{R}_d^i \mathbf{s}_k^* \right)$ and states $(s_k, ..., s_{k-L+1})$.

4.3   Estimation of the channel noise variance and the channel covariance matrix

We assumed in the previous derivation of the EM algorithm that the channel noise variance σ^2 and the channel covariance matrix $\mathbf{R}_h$ are perfectly known at the receiver. If they are not known (as it is usually the case in practical systems), they must be estimated.

The channel noise variance σ^2 can be estimated using the EM algorithm we derived previously by replacing the parameter to be estimated $\mathbf{s}$ by $\theta = (\mathbf{s}, \sigma^2)$. The function Q is then given by:

$$
\begin{aligned}
Q\left(\theta, \theta^i\right) &= \int \log P\left(\mathbf{s}, \sigma^2 | \mathbf{r}, \mathbf{h}, \hat{\mathbf{h}}\right) p\left(\mathbf{h} | \mathbf{r}, \hat{\mathbf{h}}, \mathbf{s}^i, (\sigma^i)^2\right) d\mathbf{h} \\
&= \int \log p\left(\mathbf{r} | \mathbf{s}, \sigma^2, \mathbf{h}\right) p\left(\mathbf{h} | \mathbf{r}, \hat{\mathbf{h}}, \mathbf{s}^i, (\sigma^i)^2\right) d\mathbf{h} + \log P\left(\mathbf{s}\right) - \beta \\
&= -\frac{1}{\sigma^2} \sum_{k=0}^{K+L-1} \left(|r_k - \mathbf{s}_k^T \mathbf{d}^i|^2 + \mathbf{s}_k^T \mathbf{R}_d^i \mathbf{s}_k^*\right) - (K+L)\log(\pi\sigma^2) + \log P(\mathbf{s}) - \beta \quad (18)
\end{aligned}
$$

with $\mathbf{R}_d^i$ and $\mathbf{d}^i$ being given by (15) and (16) where σ is replaced by σ^i, the estimate of σ at the ith iteration, and β is independent of θ. At the $(i+1)^{th}$ iteration, the maximization of the function Q with respect to σ leads to the noise variance estimate:

$$
(\sigma^{i+1})^2 = \frac{1}{K+L} \sum_{k=0}^{K+L-1} \left(|r_k - \mathbf{s}_k^{i^T} \mathbf{d}^i|^2 + \mathbf{s}_k^{i^T} \mathbf{R}_d^i \mathbf{s}_k^{i^*}\right). \quad (19)
$$

The equalizer needs also an estimate of the channel covariance matrix $\mathbf{R}_h$. It can be obtained by averaging (by using a moving average) over a certain number B of bursts using the channel estimates based on the TS.

4.4   Summary of the algorithm

- **Initialization**: The first estimate $\mathbf{s}^0$ of $\mathbf{s}$ is obtained by the Viterbi algorithm using the initial channel estimate $\hat{\mathbf{h}}$. The estimate of the channel noise variance $(\sigma^0)^2$ is obtained using the $T_0 + L$ training symbols: $(\sigma^0)^2 = \frac{1}{T_0} \sum_{k=t_0+L}^{t_0+T_0+L-1} \left(|r_k - \mathbf{s}_k^T \hat{\mathbf{h}}|^2\right)$. The channel covariance matrix is estimated over B bursts.

- **For each iteration index $i+1$**

  1. E-step: Compute $\mathbf{R}_d^i$ and $\mathbf{d}^i$ defined in (15) and (16), respectively, with σ^i used instead of σ.

  2. M-step: Perform a Viterbi equalizer with branch metric

$$
|r_k - \mathbf{s}_k^T \mathbf{d}^i|^2 + \mathbf{s}_k^T \mathbf{R}_d^i \mathbf{s}_k^*.
$$

  The metric minimizing sequence at the output of the Viterbi equalizer is $\mathbf{s}^{i+1}$. Compute $(\sigma^{i+1})^2$ using (19).

Remark 2: Notice that the inversions and multiplications of the matrices in the Expectation step have complexities which do not depend on K or increase linearly with K. Thus, the overall complexity of the algorithm increases linearly with K. It is worth mentioning that the complexity of the algorithm can be reduced, when K is sufficiently high, by approximating $\mathbf{S}^{i^\dagger} \mathbf{S}^i$ in (15) by $(K+L)\mathbf{I}_{L+1}$. In this case, the inversion of $\mathbf{R}_d^{i^{-1}}$ leading to $\mathbf{R}_d^i$ is done once and for all the bursts when the covariance matrices of the channel and the noise, $\mathbf{R}_h$ and $\mathbf{R}_b$, do not significantly change.

Remark 3: Notice also that our derivation of the EM algorithm is similar, from a mathematical point of view, to the one presented in [8]. However, the problem we consider is different. Indeed, in [8], Perry *et al.* propose an approximate solution, using the EM algorithm, to the problem $\hat{\mathbf{s}} = \arg \max_{\mathbf{s}} P(\mathbf{s}|\mathbf{r})$. The initialization of the EM algorithm in [8] is based on the channel statistics and does not use a training sequence (blind detection). In this paper, we propose to use the EM algorithm to give an approximate solution to the intractable problem (6) of MAP sequence equalization with an imperfect channel estimate introduced in [1]. To initialize our algorithm, we use this channel estimate $\hat{\mathbf{h}}$. As will be seen in the simulation section, when this initialization is enough reliable, the performance of our iterative equalizer approaches the performance obtained when the channel is perfectly known at the receiver.

5   Simulation results

We consider the transmission of BPSK symbols over a frequency selective channel. The channel memory length is set to $L = 4$. The different channel taps are modeled as independent zero mean complex Gaussian random variables with variance $1/(L+1)$. The channel is assumed to be invariant during a burst and to change independently from burst to burst. The channel estimate $\hat{\mathbf{h}}$ is obtained by a least mean square error estimator using a TS of length $T_0 + L$ [10]. The channel covariance matrix is estimated over $B = 20$ bursts. The channel noise variance is estimated at each iteration of the EM algorithm using (19).

5.1   Uncoded transmission

We first consider the case of an uncoded transmission, as described in Section 2. The burst length K is equal to 512. Figures 1 and 2 show the Bit Error Rate (BER) obtained using the EM algorithm from one to five iterations, with respect to $SNR = E_s/N_0$ in dB, respectively when $T_0 = 7$ and $T_0 = 12$. The training sequences are chosen randomly. The performance of the suboptimal MAP equalizer which does not take into account the channel estimation error (using $\hat{\mathbf{h}}$ as if it was the true channel $\mathbf{h}$) is similar to our iterative equalizer performance at the first iteration. Simulations show that using our iterative equalizer, which takes into account the channel estimation error, leads at the fourth iteration to a gain of 3.3dB when $T_0 = 7$ and 2.5dB when $T_0 = 12$. Simulations also show that when $T_0 = 12$, the performance of the EM algorithm at the fourth iteration approaches the performance obtained when the equalizer has perfect channel knowledge (dotted curve). However, when the TS is too short ($T_0 = 7$), there is a gap in terms of performance compared to the case of perfect channel knowledge. Thus, the TS length and then the quality of the initial channel estimate affect the asymptotic performance. This can be explained by the fact that the EM algorithm is used here to estimate discrete parameters and is then not guaranteed to converge to a local maximum [14]. Notice also that the exact MAP criterion derived in [1] cannot be implemented here in a practical way since the PDF involved in the derivation of the MAP criterion cannot be factored. The curves labeled 'Approximation [1]' show the performance obtained using the approximate rule proposed in ([1, eq(7)]). However, it leads to an improvement of only 0.2 dB compared to the one obtained at the first iteration of our iterative receiver when $T_0 = 12$ and 0.4

dB when $T_0 = 7$.

Now, we want to compare the performance of our equalizer with the performance of the iterative receiver using semiblind least squares channel estimation (LSCE) and Viterbi detection. In this receiver, at each iteration, the channel estimator uses the decisions on the transmitted symbols at the output of the Viterbi detector to improve its estimation. In Figure 3, solid curves show the performance obtained while using the proposed EM algorithm and dotted curves show the performance obtained while using the iterative receiver based on semiblind LSCE and Viterbi detection, from one to four iterations, when $T_0 = 12$. We notice that our receiver performance is slightly better at high SNR. Thus, the iterative equalizer we propose in this paper improves the global performance in terms of BER with respect to a non iterative equalizer, but it does not significantly improve the performance with respect to the iterative receiver using semiblind LSCE and Viterbi detection.
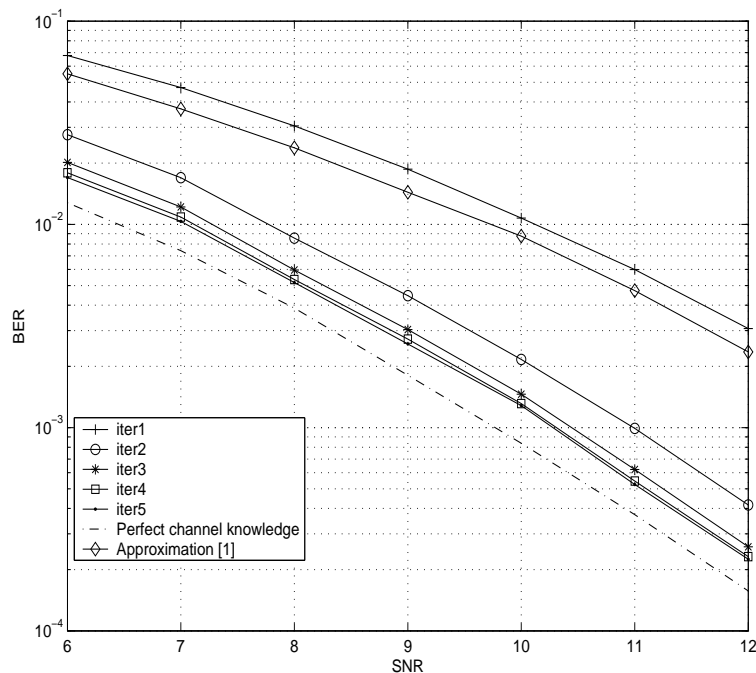

Figure 1: BER performance for one to five iterations of the EM algorithm when $T_0 = 7$ and the TS is chosen randomly.

5.2 Coded transmission

We consider now a coded data transmission system over a frequency selective channel depicted in Figure 4. The input bit sequence **b** is first encoded with a nonrecursive non-systematic convolutional encoder with rate r_c. The sequence of coded bits **c** at the output of the encoder is interleaved, mapped to the symbol alphabet $\mathcal{A}$ leading to the sequence of modulated symbols **s**. We consider at the receiver the turbo-detector consisting of a MAP equalizer and a MAP decoder which exchange extrinsic probabilities on the coded bits during the iterations [15]. The MAP decoder uses the BCJR algorithm [16]. In order to include our MAP equalizer presented in Section 4 into the turbo-detector, we need a
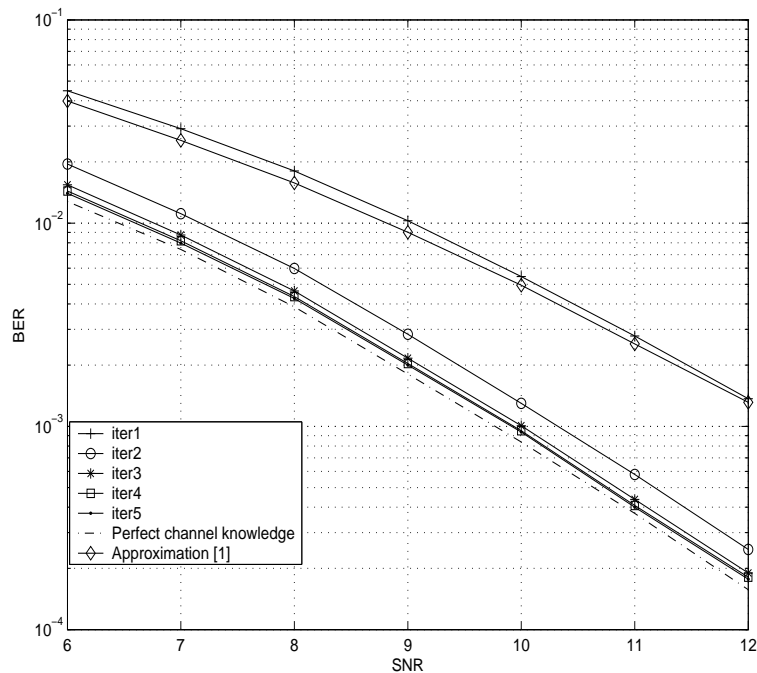
Figure 2: BER performance for one to five iterations of the EM algorithm when $T_0 = 12$ and the TS is chosen randomly.
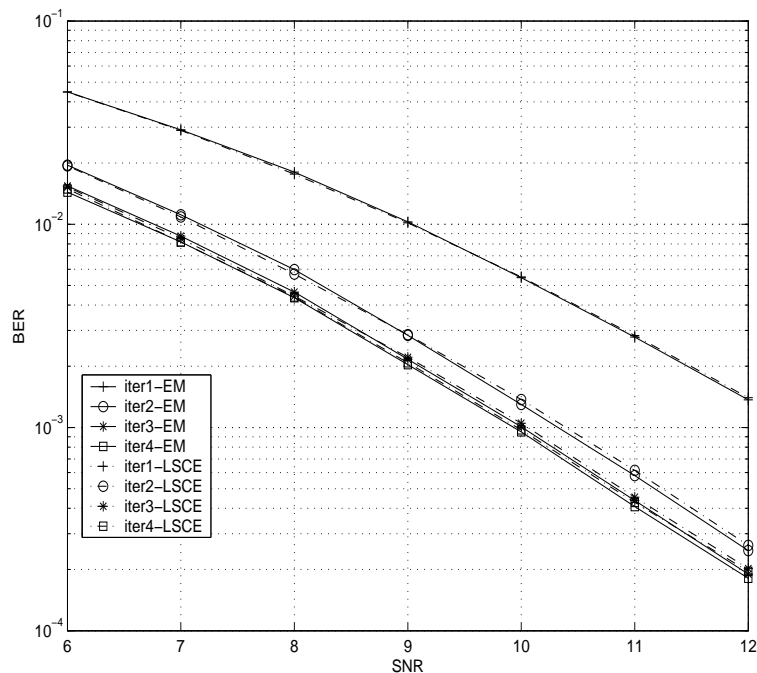

Figure 3: BER performance of the EM algorithm (solid curves) and of the iterative receiver using semiblind LSCE and Viterbi detection (dotted curves), for one to four iterations, when $T_0 = 12$ and the TS is chosen randomly.
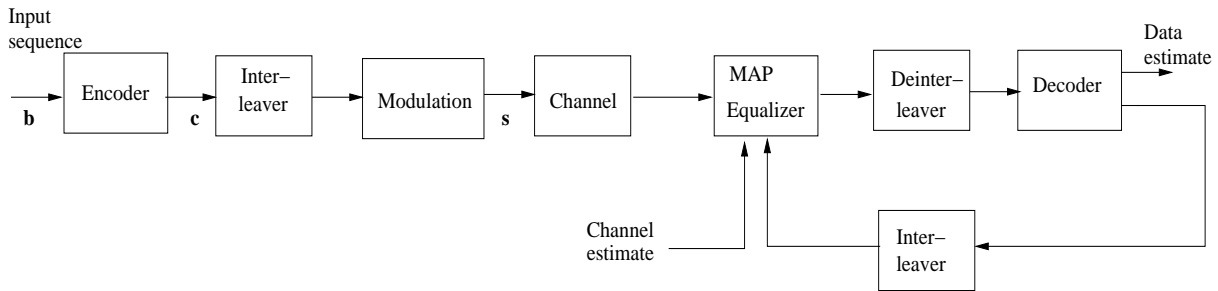
Figure 4: Coded data transmission using a turbo-detector at the receiver

soft output version of it. Since each M-step is performed by the Viterbi algorithm, we propose to use the SOVA [9]. At each iteration of the turbo-detector, an iteration of the EM algorithm is performed. The M-step is performed by a SOVA with branch metric

$$|r_k - \mathbf{s}_k^T \mathbf{d}^i|^2 + \mathbf{s}_k^T \mathbf{R}_d^i \mathbf{s}_k^* + \log P(s_k).$$

The calculations of $\mathbf{R}_d^i$ and $\mathbf{d}^i$ are performed by using (15) and (16), where the matrix $\mathbf{S}^i$ is obtained by using the estimate of the coded sequence based on the *A Posteriori* Probabilities (APPs) on the coded bits at the output of the MAP decoder. We assume that thanks to the interleaver, the APPs calculated by the decoder on the coded bits are independent. Thus, $P(s_k = s)$, $s \in \mathcal{A}$, can be computed as the product of the APPs on the coded bits provided by the decoder and corresponding to $s_k = s$.

In the simulations, the input bit sequence of length 512 (510 information bits and 2 tail bits set to 0 to end the trellis of the decoder) is encoded using the rate $r_c = 1/2$ convolutional code with 4 states and generator polynomials $(7, 5)$. In Figure 5, the curves with solid lines show the BER obtained from one to five iterations of the turbo-detector, with respect to SNR in dB for $T_0 = 10$, when the TS is chosen according to [10] (optimum TS in terms of signal to estimation error ratio). Here, $SNR = E_b/N_0$ where E_b is the energy per information bit (before coding) and $N_0 = \sigma^2$ is the channel noise variance. Notice that in this case, no performance gain is obtained at the sixth iteration. Thus, the turbo-detector converges at the fifth iteration. However, when the channel is perfectly known, only three iterations are needed by turbo-detector to converge. The curves with dotted lines show the BER obtained from one to three iterations of the turbo-detector when the equalizer knows perfectly the channel. At the convergence, the performance of the turbo-detector with imperfect channel knowledge reaches at high SNR the performance of the turbo-detector with perfect channel knowledge.

6   Conclusion

In this paper, we considered the problem of MAP sequence equalization when the channel is not perfectly known at the receiver. The MAP criterion was first derived in this case in [1] but no exact practical implementation was proposed. In this paper, we proposed an approximate solution using the EM algorithm. Each M-step is performed by the Viterbi algorithm. In order to include our equalizer into a turbo-detector, we proposed to use a soft output version of it based on the SOVA. Simulation results showed that our iterative
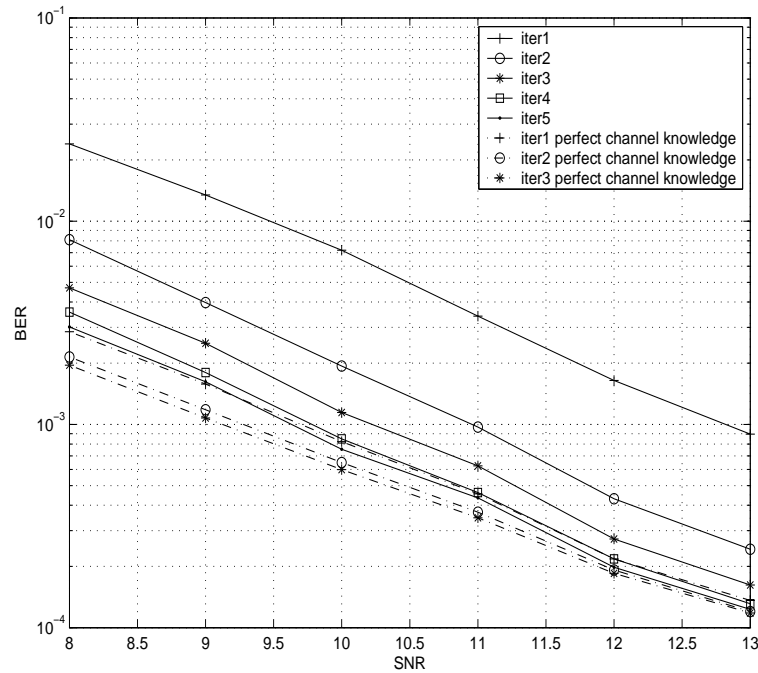
Figure 5: BER performance of the turbo-detector: for one to five iterations when $T_0 = 10$ and the TS is chosen according to [10] (solid curves), for one to three iterations when the channel is perfectly known (dotted curves).

receiver leads to a significant improvement in terms of BER performance compared to the suboptimal non iterative equalizer which does not take into account the channel estimation error. They also showed that when the initial channel estimate is enough reliable, after a few iterations, our receiver performance approaches the performance achieved when the channel is perfectly known.

References

[1] M.Tüchler and M.Mecking, "Equalization for non-ideal channel knowledge," *Conf. on Inf. Sciences and Systems*, The Johns Hopkins University, March 2003.

[2] M.Siala, "Maximum *a Posteriori* Decorrelating Discrete-Time Rake Receiver," Annals of Telecommunications, vol. 59, n° 3-4, March-April 2004.

[3] 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, December 1977.

[4] G.K. Kaleh and R. Vallet, "Joint parameter estimation and symbol detection for linear or nonlinear unknown channels," *IEEE Trans. on Communications,* vol. 42, no. 7, pp. 2406-2413, July 1994.

[5] H. Nguyen and B.C. Levy, "The Expectation-Maximization Viterbi Algorithm for Blind Adaptive Channel Equalization," *IEEE Trans. on Communications*, vol. 53, no. 10, pp. 1671-1679, October 2005.

[6] C.N.Georghiades and J.C.Han, "Sequence estimation in the presence of random parameters via the EM algorithm," *IEEE Trans. on Communications,* vol. 45, no. 3, pp. 300-308, March 1997.

[7] Y.Li, C.N.Georghiades, and G.Huang, "Iterative maximum likelihood sequence estimation for space-time coded systems," *IEEE Trans. on Communications,* vol. 49, no. 6, pp. 948-951, June 2001.

[8] R. Perry, W. Berger, and K. Buckley, "EM algorithms for sequence estimation over random ISI channels", *33-rd Asilomar Conf. on Circuits, Systems and Computers*, October 1999.

[9] J.Hagenauer and P.Höher, "A Viterbi algorithm with soft-decision outputs and its applications," *IEEE Globecom'89,*, pp.1680-1686, November 1989.

[10] S.Crozier, D.Falconer, and S.Mahmoud, "Least sum of squared errors (LSSE) channel estimation," *IEE Proceedings,* vol. 138, pp.371-378, August 1991.

[11] G.D.Forney, Jr., "Maximum-likelihood sequence estimation for digital sequences in the presence of intersymbol interference," *IEEE Trans. Inf. Theory*, vol. 18, pp. 363-378, May 1972.

[12] N. Sellami, A. Roumy, and I. Fijalkow, "The impact of both *a priori* information and channel estimation errors on the MAP equalizer performance," *IEEE Trans. on Signal Process.*, vol. 54, no. 7, pp. 2716-2724, July 2006.

[13] C. Wu, "On the convergence properties of the EM algorithm," *Annals of Stat.*, vol. 11, no. 1, pp. 95-103, 1983.

[14] H.Chen, R. Perry, and K. Buckley, "On MLSE algorithms for unknown fast time-varying channels", *IEEE Trans. on Communications,* vol. 51, no. 5, pp. 730-733, May 2003.

[15] C.Douillard, M.Jézéquel, C.Berrou, A.Picart, P.Didier, and A.Glavieux, "Iterative correction of intersymbol interference: turbo-equalization," *European Trans. Telecom.,* vol. 6, no. 5, pp. 507-511, 1995.

[16] 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.

Noura Sellami was born in Tunisia, in 1975. She received the preliminary doctorate certificate in signal processing from University of Cergy-Pontoise (UCP), France and the engineer diploma from Ecole Nationale Supérieure de l'Electronique et de ses Applications (ENSEA), France in 1999. She received the Ph.D. degree from UCP in 2002. From 1999 to 2002, she was with France Telecom R&D, Issy les Moulineaux, France. From 2003 to 2008, she was an assistant professor at l'Institut Supérieur de l'Electronique et de Communication de Sfax (ISECS), Sfax, Tunisia. Since 2008, she joined l'Ecole Nationale d'Ingénieurs de Sfax (ENIS), Sfax, Tunisia, where she is an assistant professor. She is a Member of LETI Laboratory, ENIS, Sfax, Tunisia, since 2003. Her research interests

are in the area of digital communications with special emphasis on turbo equalization, channel estimation, and optimization of the physical layer resources.

Mohamed Siala was born in 1965 in Tunisia. He received the engineer Diploma from ”Ecole Polytechnique”, Palaiseau, France, in 1988, the engineer Diploma from ”Ecole Nationale Supérieure des Télécommunications”, Paris, France, in 1990 and the PH.D. degree from the same university in 1995. From 1990 to 1992, he was with Alcatel Radio-Telephones, Colombes, France, working on the GSM physical layer. In 1995, he joined Wavecom, Issy-les-Moulineaux, France, where he worked on IOTA multicarrier modulations for terrestrial digital TV and channel estimation for the ICO project. From 1997 to 2001, he worked at France Télécom R&D, Issy-les Moulineaux, France, on the physical layer of the UMTS system. In 2001, he joined ”Ecole Supérieure des Communications de Tunis”, Tunis, Tunisia, where he is a Professor. His research interests are in the area of digital communications with special emphasis on multicarrier systems, channel estimation, modulation and coding for mobile communications.

Aline Roumy received the Engineering degree from Ecole Nationale Supérieure de l’Electronique et de ses Applications (ENSEA), Cergy, France in 1996, the Master degree in June 1997 and the Ph.D. degree in September 2000 from the University of Cergy-Pontoise, France. During 2000-2001, she has been the recipient of a French Defense DGA/DRET postdoctoral fellowship and was a research associate at Princeton University, Princeton, NJ. On November 2001, she joined INRIA, Rennes, France. Her current research and study interests include the area of statistical signal processing, coding theory and information theory.

Inès Kammoun Inès Kammoun was born in Tunisia in 1975. She received the Engineer diploma degree from Ecole Polytechnique de Tunis, Tunisia, in 1999 and the Master’s and Ph.D. degrees from Ecole Nationale Supérieure des Télécommunications (ENST), Paris, France, in 2000 and 2004, respectively. From 2004 to 2008, she was an assistant professor at ”Institut Supérieur de l’Electronique et de Communication de Sfax” (ISECS), Sfax, Tunisia. In September 2008, she joined ”Ecole Nationale d’Ingénieurs de Sfax (ENIS)”, Sfax, Tunisia, where she is an assistant professor. She is also a member of LETI laboratory of ENIS. Her main research interests are in space-time codes for noncoherent receivers, Multi-user detection and wireless communications systems.