EM Channel Estimation in a Low-cost UWB Receiver based on Energy Detection

Sami Mekki1, Jean-Luc Danger1, Benoit_MISCOPEN2, and Joseph J. Boutros3
1École Nationale Supérieure des Télécommunications, 46 rue Barrault, 75634 Cédex 13, Paris, France.
2France Télécom R&D, 28 Chemin du vieux chêne, 38243 Meylan Cédex, France.
3Texas A&M University at Qatar, Education City, PO Box 23874, Doha, Qatar.
Email: {mekki, danger}@enst.fr, benoit.miscopein, @orange-ftgroup.com, boutros@tamu.edu

Abstract—The Expectation-Maximization (EM) algorithm is studied to perform channel estimation in a low cost and high data rate impulse radio UWB receiver. The system under consideration uses a pulse position modulation with a simple analog energy detector. In order to overcome the problems inherited to high data rates, such as inter-symbol interferences, a probabilistic equalizer is used. The EM algorithm and the equalizer are embedded into the loop of an iterative channel decoder. This permits to refine both the channel parameters and the signal probability equalizer are embedded into the loop of an iterative channel decoder. This interferences, a probabilistic equalizer is used. The EM algorithm and the overcomes the problems inherent to high data rates, such as inter-symbol radio UWB receiver. The system under consideration uses a pulse

I. INTRODUCTION

The future of ultra-wideband (UWB) system is bound to the system cost at both the receiver and transmitter side. For instance, a low cost communication system could simplify the deployment of sensor or ad-hoc networks. The cheapest electronic architectures are achieved with impulse radio (IR) transmissions [20] whereas high data rates take advantage from OFDM [1] or Direct Sequence modulation schemes [9]. Among the architectures in IR UWB, the energy detector (ED) based on a simple schnottky diode and a capacitor is by far the less complex. The other alternative is the transmitted reference (TR) receiver [6], [7] or coherent detector using a Rake receiver [5], [14]. With regards to the ED receiver these electronic architectures are more complex with the necessity of analog delay lines and/or frequency mixers.

The increase of data rate in IR UWB transmission is directly dependent on the equalization process quality in a dispersive multipath channel as defined in [10]. We consider the basic IR receiver based on a low-cost ED analog front-end and a digital processor in charge of the baseband processing. The modulation is a non-linear M-ary pulse position modulation (M-PPM) which, at high data rate, is impaired with intersymbol interference (ISI) in high dispersive communication channels as those defined in [10].

To fight these impairments a new method of equalization based on a probabilistic equalizer is defined in [13]. However, this approach needs realistic channel parameters. The expectation-maximization (EM) algorithm [8] provides a numerical method for obtaining maximum likelihood of estimates that might not be available otherwise. Previous research for joint channel parameter estimation and symbol detection have been developed in [11], [18]. In [15], the EM algorithm has been investigated over random ISI channels.

In this work, we present an new method for joint channel parameters estimation and energy equalization via the EM algorithm applied to IR UWB with ED. The derived algorithm is inserted into the channel decoder loop to benefit from the iterative capacity of the decoder.

A training sequence is first exploited to get an initial estimate of the channel, then the equalizer is performed and feeds the channel decoder which in turn gives updated probabilities to the EM estimator. This loop is illustrated by a simplified schematic in Figure 1.

![Fig. 1. Simplified receiver architecture.](image-url)

This work is organized as follow: In Section II, the system model and energy distribution are specified. The considered probabilistic equalizer is described in III. Section IV presents the EM algorithm adapted to the energy detection. In section V, simulation results are presented for different interference levels and compared to the perfect channel state information (CSI) condition. Finally, a conclusion of this study is summarized in Section VI.

II. SYSTEM MODEL

We consider a pulse-based UWB transmission of a sequence of independent symbols \( c = (c_1, c_2, \ldots, c_N) \) over an additive white Gaussian noise (AWGN). We assume that the sequence \( c \) is a codeword of channel code \( C \).

As depicted in Figure 2, the encoded data is mapped into channel symbols suitable for modulation. We adopt an M-PPM modulation with \( M \) slots per symbol \( c_n \). Thus, inter-symbol and inter-slot interferences are unavoidable at high data rate transmission in a dispersive channel. The designed receiver assumes that the number of interfered symbols is \( K \), that is equivalent to \( P = (K - 1)M + 1 \) interfered slots. So, the signal at the output of the filter can be written as follows:

\[
s_n(t) = \sum_{k=0}^{K-1} x_{n-k}(t)
\]

where \( x_{n-k}(t) \) is the channel response of \( (n-k)^{th} \) transmitted symbol defined by

\[
x_{n-k}(t) = p_{n-k}(t) \otimes h(t)
\]

where \( \otimes \) denotes the convolution product, \( h(t) \) is the impulse response of the channel and

\[
p_{n-k}(t) = p(t - A_{n-k}T_{slot})
\]

is the pulse generated according to the symbol \( c_{n-k} \), where \( A_{n-k} \) takes value in \([0, 1, 2, \ldots, M - 1]\) with respect to \( c_{n-k} \) and \( T_{slot} \) is the time slot duration for an M-PPM modulation. For simulation reasons, the pulse is considered as being a Dirac \( \delta \) function, thus

\[
x_{n-k}(t) = p(t - A_{n-k}T_{slot})
\]

This assumption will not affect our
This energy can be approximated, for a process which has a band-
following expression

\[ T_0 \sigma \]

white Gaussian noise, with mean zero and variance \( \sigma^2 \).

As it is showed in [13], the energy

\[ \text{III. Energy} \]

In order to overcome the different types of interferences due to
high data rate, it is necessary to integrate an equalizer. The selected
equalizer that matches our receiver is described in [13]. Equalization
is performed according to the slot energy distribution computed by
the detector. Then, the receiver computes the probability density function

\[ p(\theta | x) \]

to get the transmitted symbol \( x_m \), this probability is given by

\[ p(\theta | x) = \sum_{x_m} \left( \prod_{i=1}^{M} p(\xi_m | x_m) \right) \int \left( p(x_m | x) \right) \]

where \( p(\xi_m | x_m) = p(\xi_m | x_m) \) defined in section II. The interested
reader should refer to [13] for details on the proof of (7). The
parameter \( B_{nm} \) defines the energy in slot \( m \) for the symbol \( n \) if the
noise is null. Let \( B \) represents the set of all possible value that \( B_{nm} \)
can take. At the receiver side, if we consider that the number on
interfered symbols does not exceed \( K \), the cardinal of \( B \) is finite. In
Table I are listed the different values of \( B \) according to \( K \), i.e. \( P \); with
a 4-PPM modulation.

\[ \text{IV. EM Algorithm for channel equalization} \]

A. EM algorithm overview

As we noticed in the previous section some specific channel
parameters must be computed to perform the equalizer. To estimate
the (CSI), the EM algorithm [3], [8] is a good candidate. It allows
build a probabilistic equalizer which could be used to feed the
decoder in the iterative decoding loop. The EM algorithm is applied
to find the maximum likelihood \( \log P(x|y, \theta) \) and it is especially effective
when the likelihood of the incomplete data is much more difficult
to maximize than the likelihood of the complete data. We denote by
incomplete data the received vector \( y \), by missing data the transmitted
vector \( x \), by complete data the couple \( (x, y) \) and by \( \theta \) the parameter
to be estimated. \( \theta \) is the channel parameter in our case.

The EM algorithm starts from an initial value of \( \theta \) and it improves
this value iteratively. This algorithm proceeds in two steps at each
iteration: the first one consists of the expectation step (E-step), and
the second one consists of the maximization step (M-step). Given a
current parameter value \( \theta \) at iteration \( i \), the EM algorithm computes
an update \( \theta^{i+1} \). The final EM estimate depends on the initial value
\( \theta \). In each iteration, the likelihood increases monotonously.

To summarize:

1) Start with \( \theta^0 \)
2) Repeat the following two steps for each iteration \( i = 1, 2, \ldots \),
   a) E-step: compute the expectation value of log-likelihood
      of complete data conditioned by observed samples and
      the current solution of \( \theta \):
      \[ Q(\theta | \theta^i) = E_x [ \log P(x, y | \theta | y, \theta) ] \]
      (8)
   b) M-step: find \( \theta^{i+1} \) that maximize the auxiliary function
      \( Q(\theta | \theta^i) \).
      \[ \theta^{i+1} = \arg \max_{\theta} Q(\theta | \theta^i) \]
      (9)

   In the case of unknown source distribution and by the means of
   Bayes’ rule and considering that \( x \) and \( \theta \) are independent, we get
   \[ p(x, y | \theta) = P(y | x, \theta) P(x | \theta) = P(y | x, \theta) \propto p(y | x, \theta) \]
   (10)
   then the new auxiliary function (8) expression is
   \[ Q(\theta | \theta^i) = E_x [ \log p(x, y | \theta) ] = \sum_{x} \log p(y | x, \theta) P(x | \theta) \]
   \[ = \sum_{x} \log p(y | x, \theta) APP (x) \]
   (11)

   where \( APP (x) = P(x | \theta) \) is the a posteriori probability of \( x \) at the
   \( i \)th iteration of the EM algorithm.

\[ \text{TABLE I} \]

Relation between the presumed number of interfering symbols \( K \) and \( | B | \) for a 4-PPM

| \( K \) | \( | B | \) |
|------|------|
| 2    | 5    |
| 3    | 9    |
| 4    | 13   |
| 5    | 15   |
| 8    | 88   |
| 4    | 424  |

We notice that \( | B | \) grows approximately in \( O(M^K) \). In the sequel, \( B_j \) refers to an element of \( B = \{ B_j \} \).
B. EM application to energy detection

As described in section II, the incomplete data is the vector of energy $\mathcal{E} = (\mathcal{E}_1, \mathcal{E}_2, \ldots, \mathcal{E}_M)$ where $\mathcal{E}_n$ is the energy per symbol that is equal to $(\mathcal{E}_{n,1}, \mathcal{E}_{n,2}, \ldots, \mathcal{E}_{n,M})$. The missing data is the vector of transmitted symbols $\mathbf{x} = (x_1, x_2, \ldots, x_N)$ and the parameter to be estimated which characterize the channel is $\theta = (\mathcal{B}, \sigma^2)$, with $\mathcal{B} = \{B\}$.

Then, the auxiliary function in energy domain is given by

$$Q(\theta | \theta') = \sum_x \log p(\mathcal{E} | \mathbf{x}, \theta) APP(\mathbf{x})$$  \hspace{1cm} (12)

Equation (12) can be decomposed into a product of probabilities by expanding the conditioned probability $p(\mathcal{E} | \mathbf{x}, \theta)$ as follows

$$p(\mathcal{E} | \mathbf{x}, \theta) = p(\mathcal{E}_1, \ldots, \mathcal{E}_M | \mathbf{x}, \theta)$$  \hspace{1cm} (13)

since the collected energy per symbol is independent from one symbol to another and it depends only on the interfering symbols, one can write

$$p(\mathcal{E} | \mathbf{x}, \theta) = \prod_{n=1}^{N} p(\mathcal{E}_n | x_{n-K+1}, \ldots, x_{n-1}, x_n, \theta) = \prod_{n=1}^{N} p(\mathcal{E}_n | B_n, \theta')$$  \hspace{1cm} (14)

the last equation comes from the unicity of the resultant energy for a given interfering symbols, i.e. $(x_{n-K+1}, \ldots, x_n)$ is equivalent to $B_n = (B_{n,1}, \ldots, B_{n,M})$. Moreover, the received energy per slot $\mathcal{E}_n$, that forms $\mathcal{E}_n$, depends on $B_n$ only throw $B_{n,m}$. Thus, equation (14) becomes

$$p(\mathcal{E} | \mathbf{x}, \theta) = \prod_{n=1}^{N} \prod_{m=1}^{M} p(\mathcal{E}_n | B_{n,m}, \theta)$$  \hspace{1cm} (15)

Applying (5) and (15) into (12), leads to the following auxiliary function

$$Q(\theta | \theta') = \sum_x \sum_{n=1}^{N} \sum_{m=1}^{M} \log p(\mathcal{E}_n | B_{n,m}, \theta) APP(\mathbf{x})$$

$$= \sum_x \sum_{n=1}^{N} \sum_{m=1}^{M} \left( \frac{1}{2\sigma^2} + \frac{L-1}{2} \log B_{n,m} - \mathcal{E}_n + \frac{B_{n,m} - \mathcal{E}_n}{2\sigma^2} + \log \frac{1}{I_{L^{-1}}(\sqrt{\frac{\mathcal{E}_n}{\sigma^2}})} \right) APP(\mathbf{x})$$  \hspace{1cm} (16)

$$= \sum_x \sum_{n=1}^{N} \sum_{m=1}^{M} \left( \frac{1}{2\sigma^2} + \frac{L-1}{2} \log B_{n,m} - \mathcal{E}_n + \frac{B_{n,m} - \mathcal{E}_n}{2\sigma^2} + \log I_{L^{-1}}(\sqrt{\frac{\mathcal{E}_n}{\sigma^2}}) \right) APP(\mathbf{x})$$  \hspace{1cm} (17)

It is noted that the missing data $[B_{n,m}]$ is also the estimated parameter. So to get rid of the parameter $\mathbf{x}$ in (17), it is necessary to rewrite $APP(\mathbf{x})$ according to $[B_{n,m}]$, we use the following approximation which has a negligible degradation on EM performance and a very low evaluation complexity [3], the a posteriori probability is conditioned on the received energy $\mathcal{E}$, since it is the only information available at the receiver:

$$\sum_x B_{n,m} APP(\mathbf{x}) = \sum_x B_{n,m} p(\mathcal{E} | \mathbf{x}, \theta)$$

$$= \sum_{x_{n-K+1}^{n}} B_{n,m} p(\mathcal{E}_n | \mathbf{x}, \theta)$$

$$= \sum_{n-K+1}^{n} B_{n,m} p(\mathcal{E}_n | x_{n-K+1}, \ldots, x_n, \theta)$$  \hspace{1cm} (18)

Equation (20) comes from the fact that $B_{n,m}$ depends only on $K$ interfering symbols in the $n^{th}$ symbol position. For each value of the interfering symbols set $(x_{n-K+1}, \ldots, x_n)$, we get a unique vector of $B_n$, so the sum over $x_{n-K+1}, \ldots, x_n$ can be replaced by the sum over the possible value that $B_n$ can take. It yields to

$$\sum_x B_{n,m} p(\mathcal{E}_n | x_{n-K+1}, \ldots, x_n \mathcal{E} | \theta') = \sum_{B_n} B_{n,m} p(\mathcal{E}_n | \theta')$$  \hspace{1cm} (21)

Proceeding as for equation (20), we get

$$\sum_{B_n} B_{n,m} p(B_{n,m} | \theta') = \sum_{B_n} B_{n,m} p(B_{n,1}, \ldots, B_{n,M} | \theta')$$

$$= \sum_{B_n} B_{n,m} p(B_{n,m} | \theta')$$  \hspace{1cm} (22)

where $B_{n,m} \in \mathcal{B}$. The updated parameters are obtained by applying (23) into (17) and deriving it with respect to $\theta$. The derivative in terms of $B_j$ of equation (23) is given by

$$\frac{\partial Q(\theta | \theta')}{\partial B_j} = \sum_{n=1}^{N} \sum_{m=1}^{M} \frac{1}{2\sigma^2} p(B_{n,m} = B_j)$$

$$+ \frac{\sqrt{\mathcal{E}_{n,m}}}{2\sigma^2} \left( \frac{\mathcal{E}_{n,m}}{\sigma} \right)^{L-1} p(B_{n,m} = B_j)$$  \hspace{1cm} (24)

Solving $\frac{\partial Q(\theta | \theta')}{\partial B_j} = 0$ with respect to $B_j$, it leads to:

$$\sqrt{B_j} \sum_{n=1}^{N} \sum_{m=1}^{M} p(B_{n,m} = B_j) = \sum_{n=1}^{N} \sum_{m=1}^{M} \frac{\mathcal{E}_{n,m}}{\sigma^2} \left( \frac{\mathcal{E}_{n,m}}{\sigma} \right)^{L-1} p(B_{n,m} = B_j)$$  \hspace{1cm} (25)

Equation (26) has no explicit solution. Some researches looked for an approximation of the non-centrality parameter of a chi-squared distribution [12], [16], [17]. We investigated a new approximation of the non-centrality parameter that has better results than those presented in the literature defined by

$$B_j^{(i+1)} = \left( \sum_{n=1}^{N} \sum_{m=1}^{M} f(\mathcal{E}_{n,m}, 2\sigma^2) p(B_{n,m} = B_j^{(i)}) \right)^2$$  \hspace{1cm} (27)

where $B_j^{(i+1)}$ is the value of $B_j$ at $(i+1)^{th}$ iteration and

$$f(\mathcal{E}_{n,m}, 2\sigma^2) = \begin{cases} \mathcal{E}_{n,m} - 2\sigma^2 & \text{if } \mathcal{E}_{n,m} > 2\sigma^2 \\ 0 & \text{if } \mathcal{E}_{n,m} < 2\sigma^2 \end{cases}$$  \hspace{1cm} (28)

where $\sigma^2$ is the updated of the noise variance at the $i^{th}$ iteration of the EM algorithm. We have not been able to show why (27) is a good approximation, but we conjecture it to be true.

To obtain the update parameter for $\sigma^2$ we proceed as for $B_j$. So, we derive the auxiliary function with respect to $\sigma^2$ and forcing the
applying (26) into the right hand side of (30), that gives
\[ \sum_{n=1}^{N} \sum_{m=1}^{M} \frac{I_{L_i}(\sqrt{E_{n,m}+B_i})}{I_{L_i-1}(\sqrt{E_{n,m}+B_i})} p(B_{n,m}=B_i) \]
This could be greater. That means that if \( B_i \) could take, so it can be replaced by the sum over \( B_i \) multiplied by the probability in slot \( n \) of the \( n \)th symbol we get \( B_i \) at the \( i \)th iteration of the EM algorithm. Such probability is defined previously by \( p(B_{n,m}=B_i) \). Equation (29) becomes then
\[ \sum_{B_j} \frac{I_{L_i}(\sqrt{E_{n,m}+B_j})}{I_{L_i-1}(\sqrt{E_{n,m}+B_j})} p(B_{n,m}=B_j) \]
The update parameter of \( \sigma^2 \) at \((i+1)\)th iteration is obtained by applying (26) into the right hand side of (30), that gives
\[ \sigma^{2(i+1)} = \frac{1}{2LMN} \sum_{n=1}^{N} \sum_{m=1}^{M} (E_{n,m} - B_{j}^{(i)}) p(B_{n,m}=B_j) \]
It should be noticed that \( \sum_{m=1}^{M} \sum_{j} p(B_{n,m}=B_j) = M \).

V. SIMULATIONS RESULTS

The simulations are computed with a bit interleaved coded modulation (BICM) [4] and a data rate of 100 Mb/s. A convolutional channel encoder at rate 1/2 with octal generator (23, 35) followed by a pseudo-random bit-inter-leaver is implemented. A 4-PPM modulation is assumed. The frame has a length of 1024 bits and the SISO decoder computes 10 iterations. We consider two hypothesis. The first one, a perfect CSI is considered: only the equalizer is implemented and the different channel parameters are given to the receiver. We simulate for different values of \( K \) which is the number of intersymbol interferences which are processed, but not the true number which could be greater. That means that if \( K \) is low, the receiver is both less complex and less effective.

In second case, the channel parameters are estimated by the mean of the EM algorithm which is combined to the SISO decoder. Only one iteration of the EM algorithm is computed per decoder iteration, so we perform a total of 10 EM iterations, since the decoder computes 10 iterations.

A. Perfect CSI condition

Perfect CSI is assumed. Figure 3 shows the BER for considered value of \( K = 2 \). The performances of the receiver can be improved in high dispersive channel; such as CM3 and CM4; if the receiver increases the value of \( K \) as shown in Figure 4 and Figure 5 for \( K = 3 \) and \( K = 4 \) respectively. In fact, the maximum excess delays for CM1, CM2, CM3 and CM4 are respectively around 50ns, 80ns, 140ns and 200ns, according to [10]. So with a 4-PPM modulation at 100Mb/s; i.e. the symbol duration \( T_s = 20 \)ns; the real number of interfered symbols for each channel model are approximately 3, 4, 7 and 10 for CM1, CM2, CM3 and CM4 respectively.

Results with \( K = 4 \) show no BER improvement. It is then preferable to stay at \( K = 3 \) because the number of energy coefficients \( B_{n,m} \) to calculate is smaller, as shown in Table I.

B. Channel estimation consideration

The EM algorithm is used and initialized by a set of training sequences. In our simulation, only 20 symbols are used as a training sequence, especially chosen to get the maximum of possible interferences. Simulations with \( K = 2 \) and 3 are depicted in Figure 6 and Figure 7 respectively. Results with \( K = 4 \), being very similar to those
with $K = 3$, are not shown in this paper.

The performance is very close to that obtained in a perfect CSI receiver. We remark that the performance of the receiver when $K = 2$; Figure 6; is better than that obtained with the perfect CSI receiver; Figure 3. This can be explained by the receiver assumption that time excess delay of the channel does not exceed $(K - 1)T_s + T_{slot}$; i.e. not more than $K = 2$ interfering symbols; but in reality, the excess delay of the channel is much longer than that. With the EM algorithm, the number of estimated energy coefficients is still 15 but the coefficients are corrected if the channel has more than 2 inter-symbol interferences. We noticed that it is not necessary to increase the complexity of the receiver with a greater value of $K$ when the data rate is 100 Mb/s, even if the channels are highly dispersive (case of CM3 and CM4).

![Fig. 6. BER for different channel models using BICM(23,35) at rate 1/2 with $K = 2$ in non perfect CSI.](image)

![Fig. 7. BER for different channel models using BICM(23,35) at rate 1/2 with $K = 3$ in non perfect CSI.](image)

VI. Conclusion

The EM algorithm has been studied to estimate the channel parameters of an M-PPM UWB communication. The parameters are composed of the noise signal and energy coefficients corresponding to $K$ inter-symbol interferences caused by high data-rate communications in dispersive channels. The channel estimation is used iteratively and jointly with a probabilistic equalization and a channel decoder. At 100 Mbits/s the EM is capable of a good estimation of parameters since the results are close to the perfect CSI implementation of the probabilistic equalizer. Moreover, the value of $K$ can be low even if the real value is greater for highly dispersive channels like CM1 and CM2. Further simulations have to be carried out at higher bit rates to study the exact value of $K$ for different communication speeds. Future works will also include the complexity refinement, as for instance the reduction of the energy coefficients to optimize the digital processing cost.

References