Blind and Semi-Blind Maximum Likelihood Techniques for Multiuser Multichannel identification.

Elisabeth de Carvalho, Luc Deneire and *Dirk T.M. Slock*[*]
EURECOM Institute, B.P. 193, 06904 Sophia Antipolis Cedex, FRANCE
e-mail : {carvalho,deneire,slock}@eurecom.fr

ABSTRACT

We investigate blind and semi-blind maximum likelihood techniques for multiuser multichannel identification. Two blind Deterministic ML methods based on cyclic prediction filters are presented [1]. The Iterative Quadratic ML (IQML) algorithm is used in [1] to solve it: this strategy does not perform well at low SNR and gives biased estimates due to the presence of noise. We propose a modification of IQML that we call DIQML to "denoise" it and explore a second strategy called Pseudo-Quadratic ML (PQML). As proposed in [2], PQML works well only at very high SNR. The solution we present here makes it work well at rather low SNR conditions and outperform DIQML. Like DIQML, PQML is proved to be consistent, asymptotically insensitive to the initialisation and globally convergent. Furthermore, it has the same performance as DML. A semi-blind extension combining these algorithms with training sequence based approaches is also studied. Simulations will illustrate the performance of the different algorithms which are found to be close to the Cramer-Rao bounds.

1 Data Model and notations

We consider a spatial division multiple access (S.D.M.A.) communication system with p emitters and a receiver constituted of an array of m_1 antennas. The signals received are oversampled by a factor m_2 w.r.t. the symbol rate, hence we have $m = m_1.m_2$ multiple channels. We assume the channels to be FIR, i.e. we assume the (vector) impulse response from source j to be of length N_j. Without loss of generality, we assume the first non-zero vector impulse response sample to occur at discrete-time zero. Let $N = \sum_{j=1}^{p} N_j$ and, w.l.o.g., $N_1 \geq N_2 \geq \cdots \geq N_p$. The discrete-time received signal can be represented as in vector form as

$$\boldsymbol{y}(k) = \sum_{j=1}^{p} \boldsymbol{H}_j A_{N_j j}(k) + \boldsymbol{v}(k) = \boldsymbol{H}\boldsymbol{A}_N(k) + \boldsymbol{v}(k) \qquad (1)$$

$$\begin{aligned}
&\boldsymbol{y}(k) = \left[y_1^H(k)\cdots y_m^H(k)\right]^H, \quad \boldsymbol{v}(k) = \left[v_1^H(k)\cdots v_m^H(k)\right]^H, \\
&\boldsymbol{h}_j(k) = \left[h_{1j}^H(k)\cdots h_{mj}^H(k)\right]^H, \boldsymbol{H}_j = \left[\boldsymbol{h}_j(0)\cdots \boldsymbol{h}_j(N_j{-}1)\right], \\
&\boldsymbol{H} = \left[\boldsymbol{H}_1\cdots \boldsymbol{H}_p\right], \quad \boldsymbol{a}(k) = \left[a_1^H(k)\cdots a_p^H(k)\right]^H, \\
&A_{nj}(k) = \left[a_j^H(k)\cdots a_j^H(k{-}n{+}1)\right]^H, \\
&\boldsymbol{A}_N(k) = \left[A_{N_1 1}^H(k)\cdots A_{N_p p}^H(k)\right]^H.
\end{aligned} \qquad (2)$$

where superscript ^{H} denotes Hermitian transpose. We consider additive temporally and spatially white Gaussian circular noise $\boldsymbol{v}(k)$ with $R_{vv}(k \Leftrightarrow i) = \mathrm{E}\left\{\boldsymbol{v}(k)\boldsymbol{v}^H(i)\right\} = \sigma_v^2 I_m \delta_{ki}$. Assume we receive M samples:

$$\boldsymbol{Y}_M(k) = \mathcal{T}_M(\boldsymbol{H})\ \boldsymbol{A}_{N+p(M-1)}(k) + \boldsymbol{V}_M(k) \qquad (3)$$

where $\boldsymbol{Y}_M(k) = \left[\boldsymbol{Y}^H(k)\cdots \boldsymbol{Y}^H(k \Leftrightarrow M+1)\right]^H$ and $\boldsymbol{V}_M(k)$ is defined similarly whereas $\mathcal{T}_M(\boldsymbol{H})$ is the multichannel multiuser convolution matrix of $\boldsymbol{H}$, with M block lines ($\mathcal{T}_M(\boldsymbol{H}) = [\mathcal{T}_M(\boldsymbol{H}_1)\ldots\mathcal{T}_M(\boldsymbol{H}_p)]$, where $\mathcal{T}_M(\boldsymbol{H}_j)$ is block Toeplitz). We shall simplify the notation in (3) with $k = 0$, and introduce the noise-free received signal as $\boldsymbol{X}$:

$$\boldsymbol{Y} = \mathcal{T}(\boldsymbol{H})\,\boldsymbol{A} + \boldsymbol{V} \quad \boldsymbol{X} = \mathcal{T}(\boldsymbol{H})\,\boldsymbol{A} \qquad (4)$$

We assume that $mM > M+N \Leftrightarrow 1$ in which case $\mathcal{T}(\boldsymbol{H})$ has more rows than columns. For obvious reasons, the column space of $\mathcal{T}(\boldsymbol{H})$ is called the signal subspace and its orthogonal complement the noise subspace.

2 Prediction-based blind Deterministic ML

2.1 Linear Prediction and Noise Subspace

Consider the problem of predicting $\boldsymbol{y}(k)$ from $\boldsymbol{Y}_L(k \Leftrightarrow 1)$, where the received signal is considered noiseless. The prediction error can be written as

$$\widetilde{\boldsymbol{y}}(k)|_{\boldsymbol{Y}_L(k-1)} = \boldsymbol{y}(k) \Leftrightarrow \widehat{\boldsymbol{y}}(k)|_{\boldsymbol{Y}_L(k-1)} = \boldsymbol{P}_L \boldsymbol{Y}_{L+1}(k) \qquad (5)$$

with $\boldsymbol{P}_L = [\boldsymbol{P}_{L,L}\cdots\boldsymbol{P}_{L,1}\ \ \boldsymbol{P}_{L,0}]$, $\boldsymbol{P}_{L,0} = I_m$. Minimising the prediction error variance leads to the following optimisation problem

$$\min_{\boldsymbol{P}_L} \boldsymbol{P}_L R_{YY} \boldsymbol{P}_L^H = \sigma_{\tilde{y},L}^2 \qquad (6)$$

hence

$$\boldsymbol{P}_L R_{YY} = \left[0\cdots 0\ \ \sigma_{\tilde{y},L}^2\right]. \qquad (7)$$

[*]The work of Luc Deneire is supported by the EC by a Marie-Curie Fellowship (TMR program) under contract No ERBFMBICT950155. The work of Elisabeth de Carvalho is supported by Laboratoires d'Electronique PHILIPS under contract Cifre 297/95. The work of Dirk Slock is supported by Laboratoires d'Electronique PHILIPS under contract LEP 95FAR008 and EURECOM Institute.

All this holds for $L \geq \underline{L} = \left\lceil \frac{Nt-p}{m-p} \right\rceil$. As a function of L, the rank profile of $\sigma^2_{\tilde{y},L}$ behaves like

$$\text{rank}\left(\sigma^2_{\tilde{y},L}\right) \begin{cases} = p & , L \geq \underline{L} \\ = m \Leftrightarrow \underline{m} \in \{p+1, .., m\} & , L = \underline{L} \Leftrightarrow 1 \\ = m & , L < \underline{L} \Leftrightarrow 1 \end{cases} \quad (8)$$

where $\underline{m} = \underline{L}(m \Leftrightarrow p) \Leftrightarrow N + p \in \{0, 1, \ldots, m \Leftrightarrow 1 \Leftrightarrow p\}$ represents the degree of singularity of $R_{YY,\underline{L}}$. For $L \geq \underline{L}$, $\tilde{\boldsymbol{y}}(k+L)|_{\boldsymbol{Y}_{L(k)}} = \boldsymbol{H}(0)\boldsymbol{a}(k+L)$. If we now consider prediction on the scalar quantities $y_i(k)$, from the former equation, we deduce that $\tilde{y}_i(k+L) = 0, i = p+1, \ldots, m$ and the $m \Leftrightarrow p$ corresponding lines of $\boldsymbol{P}_L$ are singular prediction filters (i.e. such that $\boldsymbol{P}_{L,i}\mathcal{T}(\boldsymbol{H}) = 0$). If we collect these $m \Leftrightarrow p$ singular filters and remove all dependencies between elements, we get $\overline{\boldsymbol{P}}$ of the form

$$\begin{array}{cccc|ccc} \multicolumn{4}{c}{\overbrace{\hspace{12em}}^{\underline{L}m}} & & \multicolumn{2}{c}{\overbrace{\hspace{2em}}^{\underline{m}}} \\ *..* & *..* & *..* & 10 & 0\,0\,0 & 0..0 & \\ *..* & *..* & *..* & 01 & 0\,0\,0 & 0..0 & \Big\} \underline{m} \\ \hline *..* & *..* & *..* & 00 & *\,1\,0 & 0..0 & \\ *..* & *..* & *..* & 00 & *\,*\,1 & 0..0 & \Big\} \begin{matrix} m \Leftrightarrow p \\ \Leftrightarrow \underline{m} \end{matrix} \\ & & & & \underbrace{\hspace{2em}}_{\underline{m}} & & \end{array} \quad (9)$$

where there are $N\underline{m} \Leftrightarrow p^2$ free parameters and $m \Leftrightarrow p$ 1's. One can show that $\overline{\boldsymbol{P}}$ spans the whole noise subspace, thus $\boldsymbol{H}$ can be retrieved, apart from a triangular dynamic factor (see [3]), by finding the solution of $\overline{\boldsymbol{P}}\mathcal{T}(\boldsymbol{H}) = 0$ in a least squares sense, which is equivalent to a Noise Subspace Fitting problem.

2.2 Deterministic ML

The Deterministic Maximum Likelihood (DML) method was introduced for blind channel estimation in [4, 5], and an extension to the multi-user case in [1]. In DML, both channel coefficients and input symbols are considered as deterministic quantities and are jointly estimated through the criterion:

$$\max_{\boldsymbol{A},\boldsymbol{h}} f(\boldsymbol{Y}|h) \quad \Leftrightarrow \quad \min_{\boldsymbol{A},\boldsymbol{h}} \|\boldsymbol{Y} \Leftrightarrow \mathcal{T}(\boldsymbol{H})\boldsymbol{A}\|^2 \quad (10)$$

where $f(\boldsymbol{Y}|\boldsymbol{h})$ is the probability density function. Optimising (10) w.r.t. $\boldsymbol{A}$ and replacing in (10), we get:

$$\min_h \boldsymbol{Y}^H P^{\perp}_{\mathcal{T}(\boldsymbol{H})} \boldsymbol{Y} \quad (11)$$

$P^{\perp}_{\mathcal{T}(\boldsymbol{H})}$ is the orthogonal projection on the noise subspace.

Using the linear minimal parameterisation of the Noise Subspace described here above, since $P^{\perp}_{\mathcal{T}(\boldsymbol{H})} = P_{\mathcal{T}^H(\overline{\boldsymbol{P}})}$, (11) can be written as:

$$\min_{\overline{\boldsymbol{P}}} \boldsymbol{Y}^H \mathcal{T}^H(\overline{\boldsymbol{P}})\mathcal{R}^{-1}\mathcal{T}(\overline{\boldsymbol{P}})\boldsymbol{Y} \quad (12)$$

where $\mathcal{R} = \mathcal{T}(\overline{\boldsymbol{P}})\mathcal{T}^H(\overline{\boldsymbol{P}})$. A matrix $\mathcal{Y}$ filled out with the elements of the observation vector $\boldsymbol{Y}$ can be found such that $\mathcal{T}(\overline{\boldsymbol{P}})\boldsymbol{Y} = \mathcal{Y}\overline{\boldsymbol{p}}$, where $\overline{\boldsymbol{p}}$ is a vector grouping the elements of $\overline{\boldsymbol{P}}$. Then (12) becomes (with the 'ones' and 'zeros' of $\overline{\mathbf{P}}$ remaining fixed):

$$\min_{\overline{\boldsymbol{p}}} \overline{\boldsymbol{p}}^H \mathcal{Y}^H \mathcal{R}^{-1} \mathcal{Y} \overline{\boldsymbol{p}} \quad (13)$$

One could also, as suggested in [1] introduce a vector G_N containing the free parameters of $\overline{\boldsymbol{P}}$ and $\overline{\boldsymbol{g}} = [1\ G_N^T]$, which would lead to

$$\min_{\|\overline{\boldsymbol{g}}\|=1} \overline{\boldsymbol{g}}^H \mathcal{Y}_g^H \mathcal{R}^{-1} \mathcal{Y}_g \overline{\boldsymbol{g}} \quad (14)$$

where $\mathcal{T}(\overline{\boldsymbol{P}})\boldsymbol{Y} = \mathcal{Y}_g\overline{\boldsymbol{g}}$. The constraint $\overline{\boldsymbol{g}}(0) = 1$, is equivalent to $\|\overline{\boldsymbol{g}}\| = 1$ in the minimisation of (14), hence $\hat{\boldsymbol{g}}$ is the minimal eigenvector of $\mathcal{Y}_g^H \mathcal{R}^{-1} \mathcal{Y}_g$.

2.3 Iterative Quadratic DML

The Iterative Quadratic ML algorithm (IQML) is used to solve (14) in [1]: the denominator $\mathcal{R}$, computed thanks to the previous iteration, is considered as constant and hence criterion (14) becomes quadratic. It is proved to be consistent at high SNR and requires a very good initialisation. But at low SNR conditions, it is biased because the true channel is not a stationary point of the algorithm and performs poorly.

2.4 Denoised Iterative Quadratic ML (DIQML)

We propose here a method to "denoise" the DML criterion: this denoised criterion called DIQML solved in the IQML way will be consistent.

Asymptotically in the number of data M, by the law of large numbers, (13) is equivalent to its expected value which is trace $\{P_{\mathcal{T}^H(\overline{\boldsymbol{P}})} E(\boldsymbol{Y}\boldsymbol{Y}^H)\}$. The denoising strategy consists in removing the asymptotic noise term present in $E(\boldsymbol{Y}\boldsymbol{Y}^H)$, *i.e.* $\sigma_v^2 I$; the denoised DML criterion becomes:

$$\min_{\|\overline{\boldsymbol{P}}\|=1} \text{trace}\left\{P_{\mathcal{T}^H(\overline{\boldsymbol{P}})} \left(\boldsymbol{Y}\boldsymbol{Y}^H \Leftrightarrow \sigma_v^2 I\right)\right\} \quad (15)$$

Note that this operation does not change the DML criterion solution as σ_v^2 trace $\{P_{\mathcal{T}^H(h^{\perp})}\}$ is constant. (15) is solved in the IQML way by considering $\mathcal{R}$ as constant at each iteration:

$$\min_{\overline{\boldsymbol{p}}} \overline{\boldsymbol{p}}^H \left\{\mathcal{Y}^H \mathcal{R}^+ \mathcal{Y} \Leftrightarrow \sigma_v^2 D\right\} \overline{\boldsymbol{p}} \quad (16)$$

where $\overline{\boldsymbol{p}}^H D \overline{\boldsymbol{p}} =$ trace $\{\mathcal{T}^H(\overline{\boldsymbol{P}})\mathcal{R}^+\mathcal{T}(\overline{\boldsymbol{P}})\}$. Asymptotically in the number of data, DIQML is globally convergent. Indeed, asymptotically it is equivalent to the denoised criterion:

$$\min_{\overline{\boldsymbol{p}}} \overline{\boldsymbol{p}}^H \mathcal{X}^H \mathcal{R}^+ \mathcal{X} \overline{\boldsymbol{p}} \quad (17)$$

where $\mathcal{H}\overline{\boldsymbol{p}} = \mathcal{T}(\overline{\boldsymbol{P}})\boldsymbol{X}$. When working with $\overline{\boldsymbol{g}}$, the central matrix $\overline{\boldsymbol{g}}^H \mathcal{X}_g^H \mathcal{R}^+ \mathcal{X}_g \overline{\boldsymbol{g}}$ has exactly one singularity and the solution it's the minimal eigenvector. The use of $\overline{\boldsymbol{p}}$ leads to a matrix $\overline{\boldsymbol{p}}^H \mathcal{X}^H \mathcal{R}^+ \mathcal{X} \overline{\boldsymbol{p}}$ with $(m \Leftrightarrow p)^2$ singularities, corresponding to $(m \Leftrightarrow p)^2$ ambiguities on $\overline{\boldsymbol{P}}$ if the 1's and 0's are not taken into account. Plain minimisation alleviates these indeterminacies. In practice, with large but finite M, the Hessian of (16) is indefinite: we remove a quantity λD instead of $\sigma_v^2 D$ to make it positive definite, which leads to a constrained minimisation on $\overline{\boldsymbol{p}}$ and λ. If we work with $\overline{\boldsymbol{g}}$, λ is chosen to render $\mathcal{Y}_g^H \mathcal{R}^{-1} \mathcal{Y}_g \Leftrightarrow \lambda D_g$ semi-definite (with one singularity). Hence λ is the minimal generalised eigenvalue of $\mathcal{Y}_g^H \mathcal{R}^{-1} \mathcal{Y}_g$ and D and $\overline{\boldsymbol{g}}$ is the corresponding eigenvector. Asymptotic global convergence has been proved in [6] for the single user case and extends directly here. Unfortunately, use of $\overline{\boldsymbol{g}}$ leads to merging some received signal samples and simulations show that this method yields significantly poorer performances than the use of $\overline{\boldsymbol{p}}$ (where the $m \Leftrightarrow p$ 'ones' and the 'zeros' remain fixed). In the latter case, minimisation on λ is rather tricky and asymptotic global convergence still has to be proved.

2.5 Pseudo-Quadratic ML (PQML)

The principle of PQML has been first applied to DML parameterised in terms of channel coefficients in [2]. The gradient of the DML cost function (13) may be arranged as $\mathcal{P}(\overline{\boldsymbol{p}})\overline{\boldsymbol{p}}$, where $\mathcal{P}(\overline{\boldsymbol{p}})$ is (ideally) positive semi-definite. The ML solution verifies $\mathcal{P}(\overline{\boldsymbol{p}})\overline{\boldsymbol{p}} = 0$, which is solved by the PQML strategy as follows: in a first step $\mathcal{P}(\overline{\boldsymbol{p}})$ is considered constant, so the problem becomes quadratic and as $\mathcal{P}(\overline{\boldsymbol{p}})$ is positive semi-definite the Hessian is definite positive and a solution can easily be found. This solution is used to reevaluate $\mathcal{P}(\overline{\boldsymbol{p}})$ and other iterations may be done.

The difficulty consists in finding the right $\mathcal{P}(\overline{\boldsymbol{p}})$ and especially to keep the Hessian definite positive. In our problem:

$$\mathcal{P}(\overline{\boldsymbol{p}}) = \mathcal{Y}^H \mathcal{R}^+ \mathcal{Y} \Leftrightarrow \mathcal{B}^H \mathcal{B} \tag{18}$$

$\mathcal{T}^H(\overline{\boldsymbol{P}})\boldsymbol{B} = \mathcal{B}^* \overline{\boldsymbol{p}}^*$ with $\boldsymbol{B} = \left[\mathcal{T}(\overline{\boldsymbol{P}})\mathcal{T}^H(\overline{\boldsymbol{P}})\right]^+ \mathcal{T}(\overline{\boldsymbol{P}})\boldsymbol{Y}$ (* denotes the conjugate operation). The Hessian of (13) is indefinite for finite M: the corresponding solution in [2] (where the problem is parametrised in $\boldsymbol{H}$) is to take the minimum-norm eigenvalue but this strategy does not work except for high SNR.

PQML is closely related to IQML as the first term of the central matrix in (16) and (18) are the same and $E(\mathcal{B}^H\mathcal{B}) = \sigma_v^2 D$. By analogy with IQML, we introduce an arbitrary λ and PQML becomes the following minimisation problem:

$$\min_{\|\overline{\boldsymbol{p}}\|=1,\lambda} \overline{\boldsymbol{p}}^H \left\{\mathcal{Y}^H \mathcal{R}^+ \mathcal{Y} \Leftrightarrow \lambda \mathcal{B}^H \mathcal{B}\right\} \overline{\boldsymbol{p}} \tag{19}$$

with definite positivity constraint on the Hessian.

If we work with $\overline{\boldsymbol{g}}$, λ is again chosen to render $\mathcal{Y}_g^H \mathcal{R}^{-1} \mathcal{Y}_g \Leftrightarrow \lambda \mathcal{B}_g^H \mathcal{B}_g$ semi-definite. Hence λ is the minimal generalised eigenvalue of $\mathcal{Y}_g^H \mathcal{R}^{-1} \mathcal{Y}_g$ and $\mathcal{B}_g^H \mathcal{B}_g$ and $\overline{\boldsymbol{g}}$ is the corresponding eigenvector. Asymptotic global convergence can be proved as for DIQML. The stationary points of PQML are the same as those of DML, this is why PQML has the same performance as DML. Asymptotically PQML gives the global ML minimiser (for $\overline{\boldsymbol{P}}$).

3 Semi-Blind Deterministic ML

We split the received burst according to $\boldsymbol{Y} = [\boldsymbol{Y}_{ts}^H \ \boldsymbol{Y}_b^H]^H$ where $\boldsymbol{Y}_{ts}^H = \mathcal{T}(\boldsymbol{H})A_k + \boldsymbol{V}_{ts}$ contains the observations generated by known symbols only. $\boldsymbol{Y}_b$ contains the observations generated by unknown symbols and a mixture of known and unknown symbols. Solving the DML criterion on $\boldsymbol{Y}$ would lead to joint estimation of the singular prediction filter and the channel that would have to be solved under the orthogonality constraint $\overline{\boldsymbol{P}}\mathcal{T}(\boldsymbol{H}) = 0$. We propose to use a simpler algorithm where $\overline{\boldsymbol{P}}$ is estimated in a PQML way in a first step, based on $\boldsymbol{Y}_b^H$. The channel is then estimated using the singular prediction filter relation $\overline{\boldsymbol{P}}\mathcal{T}(\boldsymbol{H}) = 0$ and the training sequence, leading to a combined criterion (note that $\boldsymbol{Y}_{ts}$ and $\boldsymbol{Y}_b$ are uncorrelated, so the linear combination of the two criteria makes sense) :

$$\min_{\boldsymbol{H}} (M \Leftrightarrow K) h^H \mathcal{T}(\overline{\boldsymbol{P}}^t) \mathcal{T}^H(\overline{\boldsymbol{P}}^t) h + K \|\boldsymbol{Y}_{ts} \Leftrightarrow \mathcal{T}(\boldsymbol{H})\boldsymbol{A}_k\|^2 \tag{20}$$

where superscript ^{t} denotes transpose of the blocks.

$$\widehat{h} = K\left((M \Leftrightarrow K).\mathcal{T}(\overline{\boldsymbol{P}}^t)\mathcal{T}^H(\overline{\boldsymbol{P}}^t) + K.\mathcal{A}_k^H \mathcal{A}_k\right)^{-1} .\mathcal{A}_k^H \boldsymbol{Y}_{ts} \tag{21}$$

where $h = \text{vec}(\boldsymbol{H})$, K is the number of known symbols and $\mathcal{A}_k h = \mathcal{T}(\boldsymbol{H})\boldsymbol{A}_k$. Factors K and $(M \Leftrightarrow K)$ are used to roughly balance the contributions of the blind and training sequence part in the criterion.

4 Simulations

We consider i.i.d. BPSK symbols, a data burst of length $M = 200$, two real channels each of length 5 and $m = 4$ sub-channels, randomly generated. The SNR is defined as $(\|\boldsymbol{H}\|^2 \sigma_a^2)/(m\sigma_v^2)$. The performance measure is the Normalised Root MSE (NRMSE) which is computed over 100 Monte Carlo runs as NRMSE $= \sqrt{\frac{1}{100}\sum_{i=1}^{100} \|\widehat{\boldsymbol{H}}^{(i)} \Leftrightarrow \boldsymbol{H}\|^2/\|\boldsymbol{H}\|^2}$ and, in the blind case, the mixing matrix U is retrieved such that, noting $\boldsymbol{H}(i) = [\boldsymbol{H}_1(i) \cdots \boldsymbol{H}_p(i)]$ and $\check{\boldsymbol{H}} = [\boldsymbol{H}(0) \cdots \boldsymbol{H}(N_1)]$, $U = \check{\boldsymbol{H}}^H \check{\boldsymbol{H}} / (\check{\boldsymbol{H}}^H \widehat{\check{\boldsymbol{H}}})$.

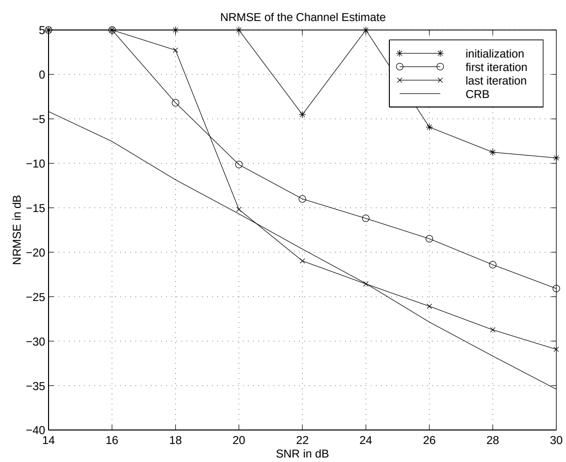

Figure 1: Performance of blind PQDML

We applied the PQDML strategy based on $\overline{\boldsymbol{p}}$, as preliminary simulations with the IQDML and PQDML gave significantly worse results when working with $\overline{\boldsymbol{g}}$. In the PQDML algorithm, we simply put λ as the generalised eigenvalue of $(\mathcal{Y}^H \mathcal{R}^{-1} \mathcal{Y}, \mathcal{B}^H \mathcal{B})$, which gave very good results as can be seen here under. Actual minimisation of the criterion w.r.t. λ under the positivity constraint of the Hessian should lead to better results at high SNR, where our simulations show that performance is not so close to the Cramer-Rao Bound.

We made five iterations in the PQML algorithm and report the performance after the first and fifth iteration. Intermediate simulations results show that the 3 first iterations should suffice.

The semi-blind algorithm has been implemented for 20 known symbols (by user), here, there is no mixing matrix to be estimated. These simulations (see figure 2) show clearly the benefit of semi-blind w.r.t. training-sequence based channel estimation.

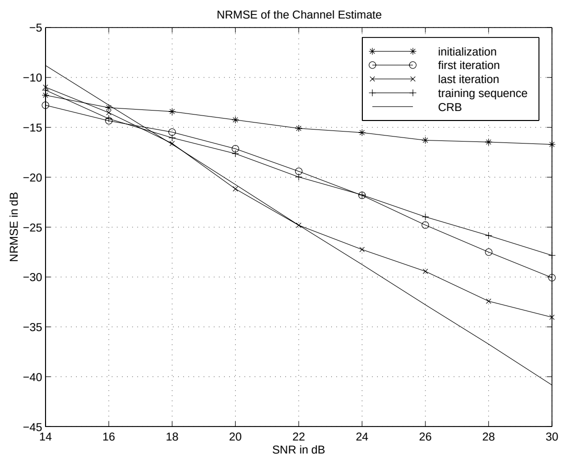

Figure 2: Performance of semi-blind PQDML

Some insight can be gained in comparing the CRB's for blind estimation (where the mixing matrix is supposed to be known) and semi-blind estimation for a few known symbols (here 10, which is insufficient to do Training-Sequence based channel estimation in our case) and 20. The closeness between the two first curves (see figure 3) show that for few known symbols, these symbols are mainly used to separate the sources (which is then perfect, opposed to blind source separation techniques) ; adding more symbols then leads to better channel estimation performance.

5 Conclusions

We have proposed several methods to solve the Blind and Semi-Blind Deterministic Maximum Likelihood criteria to estimate multiple FIR channels in a multi-user environment.

The Pseudo Quadratic Maximum Likelihood method is shown to give the global ML minimiser in $\overline{\boldsymbol{P}}$ and simulations confirm that the performances are very close to the Cramer-Rao bounds, in both the blind and semi-blind case. In the blind case, for the channel we used, we have a good performance until an SNR of 20 dB for a burst of 200 BPSK symbols. In the semi-blind case, we can go even further in the low SNRs. What the simulations show is that semi-blind approaches, in a first time, are very efficient to separate the sources and, if enough known symbols are used, lead to significantly better performances than both blind and training sequence approaches at moderate SNR. At low SNR, semi-

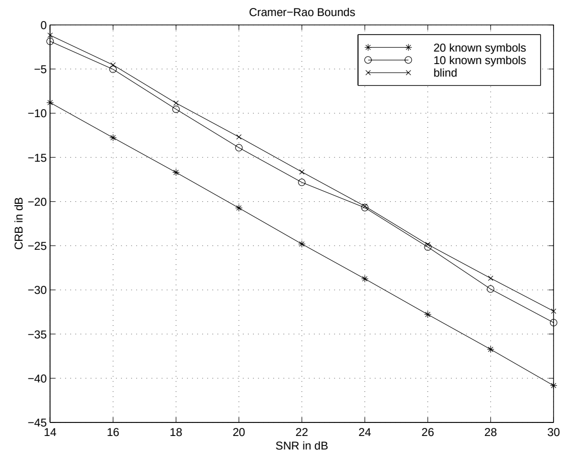

Figure 3: CRB's of blind and semi-blind DML

blind and training sequence approaches essentially yield the same results.

Further work will include refinements on the λ parameter for the minimisation of (13) and development of a global PQDML algorithm for the semi-blind approach, performing joint minimisation on $\overline{\boldsymbol{P}}$ and $\boldsymbol{H}$ using their orthogonality property .

References

[1] Dirk T.M. Slock. "Blind Joint Equalization of Multiple Synchronous Mobile Users Using Oversampling and/or Multiple Antennas". In *Twenty-Eight Asilomar Conference on Signal, Systems & Computers*, October 1994.

[2] G. Harikumar and Y. Bresler. "Analysis and Comparative Evaluation of Techniques for Multichannel Blind Deconvolution". In *Proc. Workshop Statistical Signal and Array Proc.*, pages 332–335, Corfu, June 1996.

[3] A. Gorokhov and Ph. Loubaton. "Subspace based techniques for second order blind separation of convolutive mixtures with temporally correlated sources". *IEEE Trans. on Circuits and Systems*, July 1997.

[4] D.T.M. Slock. "Blind Fractionally-Spaced Equalization, Perfect-Reconstruction Filter Banks and Multichannel Linear Prediction". In *Proc. ICASSP Conf.*, Adelaide, Australia, April 1994.

[5] D.T.M. Slock. "Subspace Techniques in Blind Mobile Radio Channel Identification and Equalization Using Fractional Spacing and/or Multiple Antennas". In *"Proc. 3rd International Workshop on SVD and Signal Processing"*, Leuven, August 1994.

[6] Jahouar Ayadi, Elisabeth de Carvalho, and Dirk Slock. "Blind and Semi-Blind Maximum Likelihood Methods for FIR multichannel identification". In *ICASSP*, Seattle, U.S.A., May 1998.