An empirical likelihood method for data aided channel identification in unknown noise field

Frédéric Pascal, Jean-Pierre Barbot, Hugo Harari-Kermadec, Ricardo Suyama, Pascal Larzabal

AN EMPIRICAL LIKELIHOOD METHOD FOR DATA AIDED CHANNEL IDENTIFICATION IN UNKNOWN NOISE FIELD

Frédéric Pascal[1], *Jean-Pierre Barbot*[2], *Hugo Harari-Kermadec*[3], *Ricardo Suyama*[4], *and Pascal Larzabal*[2]

[1] SONDRA, Supelec, Plateau du Moulon, 3 rue Joliot-Curie, F-91192 Gif-sur-Yvette Cedex, France
[2] SATIE, ENS Cachan, CNRS, UniverSud, 61, av President Wilson, F-94230 Cachan, France
[3] CREST-LS and University Paris-Dauphine, France,
[4] DSPCOM, DECOM,FEEC, State University of Campinas, C.P. 6101, Campinas, SP, Cep 13083-970, Brazil

email: frederic.pascal@supelec.fr, barbot@satie.ens-cachan.fr, harari@ensae.fr, rsuyama@decom.fee.unicamp.br, larzabal@satie.ens-cachan.fr

ABSTRACT

In this work we propose a different approach to the problem of estimating a transfer matrix with data aided, which can be applied to SISO and SIMO channels, either baud-rate or fractionally sampled signals. The approach is based on the Empirical Likelihood method [1, 2], a flexible semi-parametric estimation method, which can easily integrate in the estimation procedure some prior informations on the structure of the parameter of interest. Moreover, this improved estimation method does not assume any model for the data distribution. The contributions of this paper is twofold: first, we introduce the Empirical Likelihood method in a general context, i.e. without any prior informations and then, we derive closed-form expressions of estimators for the transfer matrix. All results are presented under Gaussian assumptions and under a mixture of Gaussian and Student-t distributions. This allows to show the robustness of the proposed method.

1. INTRODUCTION

The main objective of a communication system is to allow information to be properly exchanged between a transmitter and a receiver that are interconnected by a channel. As a rule, this communication channel will introduce distortions (intersymbol interference, noise, etc.) that are usually compensated for using an equalizer, whose design often needs an accurate estimator of the channel [3].

Traditionally, a training sequence is used to aid the channel estimation task. This emitted known sequence makes the identification of the channel feasible since both the input and the output signals are known during the transmission of the training sequence. Note that most of the propagation channel estimators are derived under the hypothesis of an additive Gaussian noise model. With the resultant propagation channel coefficients, several linear and nonlinear methods can be used to estimate the emitted symbols [3].

In this work we propose a different approach to the problem, which can be applied to SISO and SIMO channels, either baud-rate or fractionally sampled signals. The approach is based on the empirical likelihood [1, 2], a flexible semi-parametric estimation method, which can easily incorporate a priori knowledge about the problem at hand.

In order to expose the proposed technique, this paper is organized as follows. Section 2 presents the problem of interest. In Section 3 some background on the EL procedure is provided. The section 4 details the introduction of prior on the transfer matrix. Section 5 presents the simulation results, followed by the concluding remarks in Section 6.

2. PROBLEM FORMULATION

2.1 Mathematical notation

In the following, boldface letters (respectively capital letters) denote column vectors (resp. matrices), H denotes the conjugate transpose operator, T denotes the transpose operator, $E[\cdot]$ stands for the expectation of a random variable, $E_P[\cdot]$ is the expectation under the data probability P. $\mathbb{C}$ (resp. $\mathbb{R}$) denotes the set of complex (resp. real) numbers, while for any integer p, $\mathbb{C}^p$ (resp. $\mathbb{R}^p$) represents the set of p-vectors with complex (resp. real) elements. For $z \in \mathbb{C}$, we write $\mathscr{R}e(z)$ and $\mathscr{I}m(z)$ its real and imaginary parts.

2.2 Signal Model

The proposed signal model is the same as in [4]. Consider a SIMO channel, modeled by a set of L finite response filters (FIR) of length $M+1$, each one composed of a set of taps $h_k^{(i)}$, $k = 0,\dots,M$, $i = 0,\dots,L-1$ and a transmitted signal composed of a finite sequence of symbols s_k. The received signal $x_k^{(i)}$ from the i-th subchannel is then given by

$$x_k^{(i)} = \sum_{l=0}^{M} h_k^{(i)} s_{l-k} + n_k^{(i)}, \tag{1}$$

where $n_k^{(i)}$ denotes an additive white Gaussian noise of unknown variance. The transmitted sequence s_k is assumed to be i.i.d, digitally modulated signal, and the channel impulse response is supposed time-invariant during the observation record. The data aided channel estimation task consists of determining the set of coefficients $h_k^{(i)}$ based on a set of observed samples $x_k^{(i)}$ and with the knowledge of the emitted signal.

Let $\mathbf{x}_k^{(i)} = [x_k^{(i)}, x_{k-1}^{(i)}, \cdots, x_{k-N+1}^{(i)}]^T$ represent a vector containing N samples of the i-th subchannel output. Denoting by $\mathbf{s}_k = [s_k, s_{k-1}, \cdots, s_{k-M-N}]^T$ the vector containing $M+N$ transmitted symbols, it is possible to express $\mathbf{x}_k^{(i)}$ as follows:

$$\mathbf{x}_k^{(i)} = \mathbf{H}^{(i)} \mathbf{s}_k + \mathbf{n}_k^{(i)}, \tag{2}$$

where $\mathbf{n}(k)$ represents the vector of additive Gaussian noise, and

$$\mathbf{H}^{(i)} = \begin{bmatrix} h_0^{(i)} & h_1^{(i)} & \cdots & h_M^{(i)} & 0 & \cdots & 0 \\ 0 & h_0^{(i)} & h_1^{(i)} & \cdots & h_M^{(i)} & \cdots & 0 \\ \vdots & & & & & & \vdots \\ 0 & \cdots & 0 & h_0^{(i)} & h_1^{(i)} & \cdots & h_M^{(i)} \end{bmatrix} \quad (3)$$

denotes the convolution matrix related to the i-th subchannel. Hence, stacking the received vectors $\mathbf{x}_k^{(i)}$ into a single vector $\mathbf{x}_k$, yields the following received signal model:

$$\underbrace{\begin{bmatrix} \mathbf{x}_k^{(0)} \\ \vdots \\ \mathbf{x}_k^{(L-1)} \end{bmatrix}}_{\mathbf{x}_k} = \underbrace{\begin{bmatrix} \mathbf{H}^{(0)} \\ \vdots \\ \mathbf{H}^{(L-1)} \end{bmatrix}}_{\mathbf{H}} \mathbf{s}_k + \underbrace{\begin{bmatrix} \mathbf{n}_k^{(0)} \\ \vdots \\ \mathbf{n}_k^{(L-1)} \end{bmatrix}}_{\mathbf{n}_k} \quad (4)$$

The following section is devoted to the estimation of the transfer matrix $\mathbf{H}$. First, we introduce an improved estimation procedure, the Empirical Likelihood and then, we estimate $\mathbf{H}$ with prior informations on its structure since $\mathbf{H}$ is a Sylvester matrix.

3. EMPIRICAL LIKELIHOOD

The Empirical Likelihood method is a recent semi-parametric method [5] designed for estimation problems in which the parameter of interest $\mathbf{H}$ is defined as the solution of an estimating equation.

The covariance matrix estimation problem, that we consider in this paper, can be formulated in such a way:

$$E_{P_0}\left[\mathbf{x}\mathbf{s}^H - \mathbf{H}\right] = \mathbf{0}. \quad (5)$$

where $\mathbf{0}$ denotes the null vector with appropriate dimension ($n = NL \times [N+M-1]$ here) since the signal $\mathbf{s}$ is a random vector containing 1 or -1 and such that $E[\mathbf{s}\mathbf{s}^H] = \mathbf{I}$. Let us recall that NL is the dimension of the observations $\mathbf{x}$ and $N+M-1$ the dimension of the signal $\mathbf{s}$. Notice that the expectation is taken under the true Probability Density Function P_0 of $\mathbf{x}$ which is assumed to be unknown. This is an important feature of the EL procedure which does not require any assumptions on the data distribution. In most of the time, the additive noise is assumed to be Gaussian.

In this paper, we only use the EL method to estimate the transfer matrix $\mathbf{H}$. Therefore, we will restrict ourselves to built the corresponding Empirical Likelihood and to derive its Maximum Empirical Likelihood (MEL) estimator.

3.1 Likelihood

In this section, we introduce EL methodology from the classical likelihood context. For that purpose, we consider the family of multinomials G charging the data set as if it was a parametric model for the data. Notice that this parametric model assumption will never be made in our method, we just made it to interpret EL as a classical likelihood. All the details and technical arguments are given in Owen's book [5].

We define, for G and $\mathbf{H}$ verifying the moment condition $E_G\left[\mathbf{x}\mathbf{s}^H - \mathbf{H}\right] = \mathbf{0}$, the distribution

$$G(\mathbf{x}) = \begin{cases} q_k & \text{if } \exists k,\ \mathbf{x} = \mathbf{x}_k \\ 0 & \text{otherwise}, \end{cases} \quad (6)$$

or equivalently,

$$G(\mathbf{x}) = \sum_{k=1}^{K} q_k\, \delta_{\mathbf{x}_k}(\mathbf{x}), \quad (7)$$

where $\delta_{\mathbf{x}}$ stands for the Dirac measure at element $\mathbf{x}$, $0 < q_k < 1$ and $\sum_{k=1}^{K} q_k = 1$.

The corresponding likelihood, as if the data were distributed according to G, is called Empirical Likelihood:

$$\begin{aligned} EL(\mathbf{H}) &= EL(\mathbf{x}_1, \cdots, \mathbf{x}_K, \mathbf{s}_1, \cdots, \mathbf{s}_K, \mathbf{H}) \\ &= \sup_{(q_k)} \left\{ \prod_{k=1}^{K} q_k \,\middle|\, \sum_{k=1}^{K} q_k \left(\mathbf{H} - \mathbf{x}_k\mathbf{s}_k^H\right) = \mathbf{0},\ \sum_{k=1}^{K} q_k = 1 \right\}. \end{aligned} \quad (8)$$

The main technical difficulty, evaluating the empirical likelihood $EL(\mathbf{H})$ at any given $\mathbf{H}$, is resolved by a Lagrangian method and $EL(\mathbf{H})$ can be written in an easier form:

$$-2\log\left(EL(\mathbf{H})\right) = \inf_{\lambda} \left\{ 2\sum_{k=1}^{K} \log\left(K\left(1 + \lambda^\top \mathrm{vec}\left(\mathbf{H} - \mathbf{x}_k\mathbf{s}_k^H\right)\right)\right)\right\}, \quad (9)$$

because the optimal weights write

$$q_k^* = \frac{1}{K}\left(1 + \lambda^{*\top} \mathrm{vec}\left(\mathbf{H} - \mathbf{x}_k\mathbf{s}_k^H\right)\right)^{-1}, \quad (10)$$

where λ^* is the optimal Lagrange multiplier and depends on $\mathbf{H}$ and where the operator $\mathrm{vec}(\cdot)$ reshapes a $m \times n$ matrix elements into a mn column vector .

3.2 Maximum Empirical Likelihood without prior information

Now, $\widehat{\mathbf{H}}_{EL}$ is defined as the argsup of $EL(\mathbf{H})$ which is the arginf of $-2\log\left(EL(\mathbf{H})\right)$ since $-2\log(.)$ is a decreasing function :

$$\widehat{\mathbf{H}}_{EL} = \arg\inf_{\mathbf{H}} \left(\inf_{\lambda} \left\{ 2\sum_{k=1}^{K} \log\left(K\left(1 + \lambda^\top \mathrm{vec}\left(\mathbf{H} - \mathbf{x}_k\mathbf{s}_k^H\right)\right)\right)\right\}\right). \quad (11)$$

If no restriction is assumed on the structure of $\mathbf{H}$, the Maximal Empirical Likelihood (MEL) estimator is given by the following expression

$$\widehat{\mathbf{H}}_{EL1} = \overline{\mathbf{x}\mathbf{s}^H} = \frac{1}{K}\sum_{k=1}^{K} \mathbf{x}_k\mathbf{s}_k^H \quad (12)$$

where the notation $\overline{\mathbf{a}}$ is used for the empirical mean of the vector $\mathbf{a}$: $\frac{1}{K}\sum_{k=1}^{K} \mathbf{a}_k$.

Notice that $\widehat{\mathbf{H}}_{EL1}$ is equal to the Maximum Likelihood (ML) estimator under Gaussian assumption. This is the purpose of the following proposition:

Proposition 3.1
Let $(\mathbf{x}_1,\ldots,\mathbf{x}_K)$ an i.i.d. data set in $\mathbb{R}^p$, with common distribution P_0 and expectation $\theta_0 \in \mathbb{R}^p$ and finite variance-covariance matrix. Then the Maximal Empirical Likelihood estimator of the expectation is given by

$$\hat{\theta}_{MEL} = \overline{\mathbf{x}}\,. \tag{13}$$

where $\overline{\mathbf{x}}$ denotes the empirical mean, $\overline{\mathbf{x}} = \dfrac{1}{K}\sum_{k=1}^{K}\mathbf{x}_k\,.$

Remark 3.1
This result is interesting since it provides an estimator of the mean without any assumptions on the data distribution and it corresponds to the Maximum Likelihood estimator under Gaussian assumptions (and many others classical distributions).

Another important feature of the EL method is that prior informations can be taken into account into the estimation procedure. This is one of the two contribution of this paper.

3.3 Additional prior information

This explicit expression given by equation (12) for the estimator can be obtained even when some restrictions are made on the covariance matrix structure. The particular case of Sylvester matrix will be considered in section 4. The estimation scheme can take into account some prior information by introducing an additional equation to equation (5) as follows:

$$E_{P_0}\left[\begin{array}{c}\mathrm{vec}(\mathbf{H}-\mathbf{x}\mathbf{s}^H)\\ \mathbf{c}^{(prior)}\end{array}\right] = \begin{pmatrix}\mathbf{0}_{p^2}\\ \mathbf{0}_s\end{pmatrix}. \tag{14}$$

where $\mathbf{c}^{(prior)}$ is a vector resulting in a data transformation which reflects the priors information and where p^2 (resp. s) is the size of $\mathrm{vec}(\mathbf{H}-\mathbf{x}\mathbf{s}^H)$ (resp. $\mathbf{c}^{(prior)}$)

By rewriting equation (14) as

$$E_{P_0}\left[\begin{array}{c}\mathbf{v}\\ \mathbf{w}\end{array}\right] = \begin{pmatrix}\mathrm{vec}(\mathbf{H})\\ \mathbf{0}\end{pmatrix}, \tag{15}$$

where $\mathbf{v} = \mathrm{vec}(\mathbf{x}\mathbf{s}^H)$, $\mathbf{w} = \mathbf{c}^{(prior)}$ and by setting, for the covariance matrix of the vector $\begin{pmatrix}\mathbf{v}\\ \mathbf{w}\end{pmatrix}$

$$\mathrm{Var}\begin{pmatrix}\mathbf{v}\\ \mathbf{w}\end{pmatrix} = \begin{pmatrix}V_{11} & V_{12}\\ V_{21} & V_{22}\end{pmatrix}, \tag{16}$$

one obtains the following closed-form expression for the MEL estimator:

$$\hat{\theta}_{MEL} = \overline{\mathbf{v}} - V_{12}V_{22}^{-1}\overline{\mathbf{w}}\,. \tag{17}$$

This closed-form can only be obtained when priors information can be written in terms of equation (14). This result has been proved in [5], page 52. In practice, the empirical versions of V_{12} and V_{22} allow to obtain an approximation of the MEL, without significant computational cost.

Notice that when there is no prior information, i.e. $\mathbf{w} = \mathbf{0}$, one obtains $\widehat{\mathbf{H}}_{EL1}$. This is also the case when $\mathbf{w}$ is uncorrelated with $\mathbf{v}$ because V_{12} is the null matrix.

4. APPLICATION TO SYLVESTER MATRIX ESTIMATION

In this section, we use prior information on the transfer matrix structure, defined in equation (4). This section is divided in three steps:

- It is first assumed that $\mathbf{H}$ is a real matrix.
- Then, the full Sylvester structure of $\mathbf{H}$ is contained in $\mathbf{c}^{(prior)}$.

Notice that there already exist methods for structured covariance matrix estimation in which the Toeplitz case is treated, see e.g. [6, 7], but these methods are based on a parametric model, usually the Gaussian one.

4.1 First prior information: H is real

A first step is to assume that $\mathbf{H}$ has real valued elements. Therefore, the estimating equation is modified to take the structure into account. For that purpose, we define the function $\mathbf{m}_2$ as

$$\mathbf{m}_2(\mathbf{x},\mathbf{s},\mathbf{H}) = \begin{pmatrix}\mathscr{R}e\left(x_{j1}s_{j2}^H - h_{j1j2}\right)_{1\le j1\le NL,1\le j2\le M+L-1}\\ \mathscr{I}m\left(x_{j1}s_{j2}^H - h_{j1j2}\right)_{1\le j1\le NL,1\le j2\le M+L-1}\end{pmatrix}. \tag{18}$$

This leads to a new estimator $\widehat{\mathbf{H}}_{EL2}$ which integrates the constraint on the real transfer matrix $\mathbf{H}$. This writes

$$\widehat{\mathbf{H}}_{EL2} = \arg\inf_{(\mathbf{H},\lambda)}\left\{\sum_{k=1}^{K}\log\left(1+\lambda^\top\mathbf{m}_2(\mathbf{x}_k,\mathbf{s}_k,\mathbf{H})\right)\right\}. \tag{19}$$

4.2 Second prior information: H shares a Sylvester structure with $L\times M$ non null elements

In this subsection, the unknown parameters of $\mathbf{H}$ are only $L\times M$ real scalars: for $i=1,\ldots,L$, $h_1^{(i)},\ldots,h_M^{(i)}$. To estimate $\mathbf{H}$ by taking into account the prior informations, we will focus on each Sylvester matrix $\mathbf{H}^{(i)}$.

$$\mathbf{m}_3(\mathbf{x}^{(i)},\mathbf{s},\mathbf{H}^{(i)}) \;=\; \left(\left(x_{j1}^{(i)}s_{j1+j2}^H - h_{j2}^{(i)}\right)_{1\le j1\le N,\,0\le j2\le M}\right) \tag{20}$$

This leads to the last estimator $\widehat{\mathbf{H}}_{EL3}$ of $\mathbf{H}$, defined by

$$\widehat{\mathbf{H}}_{EL3} = \left[\begin{array}{c}\widehat{\mathbf{H}}^{(0)}\\ \vdots\\ \widehat{\mathbf{H}}^{(L-1)}\end{array}\right] \tag{21}$$

where

$$\widehat{\mathbf{H}}^{(i)} = \arg\inf_{(\mathbf{H}^{(i)},\lambda)}\left\{\sum_{k=1}^{K}\log\left(1+\lambda^\top\mathbf{m}_3(\mathbf{x}_k^{(i)},\mathbf{s}_k,\mathbf{H}^{(i)})\right)\right\}. \tag{22}$$

We can rewrite the constraint in terms of expectations in order to obtain an explicit form of the estimator by means of equation (17). For simplicity purpose, we give the constraints for $M = 2$ and $N = 2$. Let

$$\mathbf{v} = \mathscr{R}e\left(x_1 s_1^H, x_1 s_2^H\right)^\top$$

$$\mathbf{w} = \left(\mathscr{I}m\left(x_1 s_1^H, x_1 s_2^H\right), x_1 s_1^H - x_2 s_2^H, x_1 s_2^H - x_2 s_3^H, x_1 s_3^H, x_2 s_1^H\right)^\top \quad (23)$$

Let us explain constraints containing in the vector $\mathbf{w}$:

- The first elements allows to specify that h_1 and h_2 are real scalars, thus their imaginary part is null.
- The two subtractions mean that the two diagonals of $\mathbf{H}$ contains the same element, and thus the difference is null.
- Finally, last elements are the null elements of the transfer matrix.

Equation (17) gives estimators for the first line of $\mathbf{N}$:

$$(\widehat{h}_1, \widehat{h}_2)^\top = \overline{\mathbf{v}} - V_{12} V_{22}^{-1} \overline{\mathbf{w}} \quad (24)$$

The estimator $\widehat{\mathbf{H}}^{(i)}$ writes then

$$\widehat{\mathbf{H}}^{(i)} = \begin{bmatrix} \widehat{h}_1 & \widehat{h}_2 & 0 \\ 0 & \widehat{h}_1 & \widehat{h}_2 \end{bmatrix} \quad (25)$$

Remark 4.1
One can give a general expression for all these estimators:

$$\widehat{\mathbf{H}}_{ELj} = \sum_{k=1}^{K} q_k^*(j)\, \mathbf{x}_k \mathbf{s}_k^H,$$

where the $q_k^*(j)$ depend on the constraints. For example, in the case of no constraint (i.e. $\widehat{\mathbf{H}}_{EL1}$), the $q_k^*(1)$ are all equal to K^{-1}. This corresponds to the ML estimator under Gaussian assumptions.

These theoretical estimators of $\mathbf{H}$ will be studied in the section 5 thanks to simulations on their Mean Square Error (MSE) under Gaussian and non-Gaussian assumptions.

5. SIMULATIONS

In order to enlighten results provided in sections 3 and 4, some simulation results are presented. We focus on the problem of transfer matrix estimation in the case of an additive Gaussian or non-Gaussian noise.

In order to compare all estimators, we will plot the Mean Square Error (MSE) in different realistic situations. The MSE used in this section is the following criterion:

$$MSE(\widehat{\mathbf{H}}, \mathbf{H}) = E\left[\frac{\|\widehat{\mathbf{H}} - \mathbf{H}\|}{\|\mathbf{H}\|}\right],$$

where $\|.\|$ stands for the Frobenius norm and $\widehat{\mathbf{H}}$ denotes the studied estimator of $\mathbf{H}$.

The transfer matrix $\mathbf{H}$ which has to be estimated shares a Sylvester structure and is defined as follows:

$$\mathbf{H} = \begin{bmatrix} \mathbf{H}^{(0)} \\ \vdots \\ \mathbf{H}^{(L-1)} \end{bmatrix}, \quad (26)$$

where

$$\mathbf{H}^{(i)} = \begin{bmatrix} h_0^{(i)} & h_1^{(i)} & \cdots & h_M^{(i)} & 0 & \cdots & 0 \\ 0 & h_0^{(i)} & h_1^{(i)} & \cdots & h_M^{(i)} & \cdots & 0 \\ \vdots & & & & & & \vdots \\ 0 & \cdots & 0 & h_0^{(i)} & h_1^{(i)} & \cdots & h_M^{(i)} \end{bmatrix}. \quad (27)$$

In the simulations, the number L of response filters is $L = 4$, the length $M+1$ of each response filter is $M+1 = 5$ and the number of samples is $N = 10$. Thus, the matrix $\mathbf{H}$ has $N \times L$ lines and $M+N-1$ columns. Moreover, the number K of $\mathbf{x}_k$ and $\mathbf{s}_k$ is $K = 200$. The transfer matrix $\mathbf{H}$ to be estimated is defined thanks to the following response filter:

- $h^{(1)} = [-0.049; 0.482; -0.556; 1; -0.171]$,
- $h^{(2)} = [0.443; 1; 0.921; 0.189; -0.087]$,
- $h^{(3)} = [-0.211; -0.199; 1; -0.284; 0.136]$,
- $h^{(4)} = [0.417; 1; 0.873; 0.285; -0.049]$.

To evaluate performance of our method, we will compare the three following estimators of $\mathbf{H}$:

- The well-known Sample Covariance Matrix which corresponds to the ML estimator under Gaussian assumptions and defined as follows
$$\widehat{\mathbf{H}}_{ML} = \frac{1}{K}\sum_{k=1}^{K} \mathbf{x}_k \mathbf{s}_k^H.$$
$\widehat{\mathbf{H}}_{ML}$ is used as a benchmark but it is not appropriate to our problem since it does not take into account the structure of the real covariance matrix. Notice that this ML estimator is the same as $\widehat{\mathbf{H}}_{EL1}$ which is the MEL estimator without any prior informations.
- Then, we plot $\widehat{\mathbf{H}}_{EL2}$ which is the MEL estimator in the case of a real transfer matrix $\mathbf{H}$. $\widehat{\mathbf{H}}_{EL2}$ is defined by equation (19).
- Finally, we plot $\widehat{\mathbf{H}}_{EL3}$, the MEL estimator which uses all prior information on the structure. $\widehat{\mathbf{H}}_{EL3}$ is defined by equation (25).

Concerning the EL method, notations of section 4 are still valid: $\widehat{\mathbf{H}}_{EL1}$, $\widehat{\mathbf{H}}_{EL2}$ and $\widehat{\mathbf{H}}_{EL3}$. Notice that both $\widehat{\mathbf{H}}_{EL2}$ and $\widehat{\mathbf{H}}_{EL3}$ are analytically derived thanks to equation 17. Matrices V_{12} and V_{22} are replaced by their empirical versions. Finally, for $\widehat{\mathbf{H}}_{EL3}$, $\mathbf{v}$ and $\mathbf{w}$ are given by equation 23 and for $\widehat{\mathbf{H}}_{EL2}$, $\mathbf{v}$ and $\mathbf{w}$ are straightforward.

We first studied the behavior of the proposed method in a Gaussian context and then in a non-Gaussian context.

5.1 Gaussian distribution

Figure 1 shows the MSE of each estimator versus the Signal-to-Noise Ratio under Gaussian assumptions, i.e. the additive noise $\mathbf{b}_k$ is Gaussian distributed.

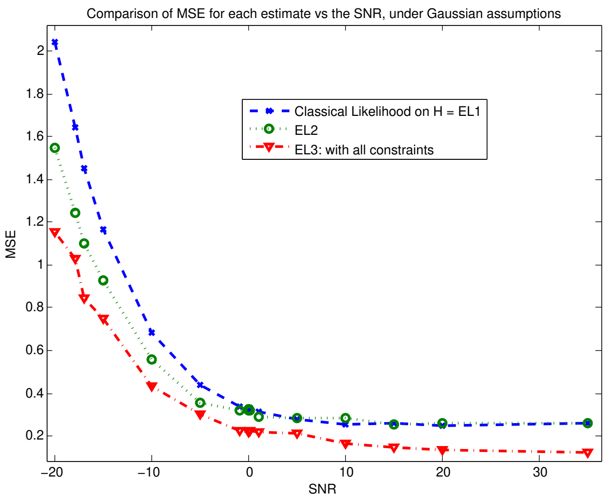

Figure 1: MSE against the Signal-to-Noise Ratio, under Gaussian assumptions, for $L = 4$, $M = 5$, $N = 10$ and $SNR = -17dB$.

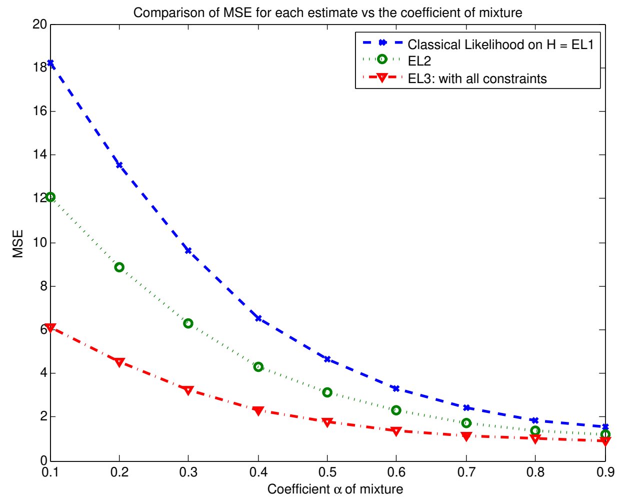

Figure 2: MSE against α, for a mixture of Student-t and Gaussian distributions, for $L = 4$, $M = 5$ and $N = 10$.

5.2 Mixture of Student-t and Gaussian distributions

For $k = 1, \ldots, K$,

$$\mathbf{b}_k \sim \frac{(1-\alpha)\mathscr{C}\mathscr{T} + \alpha \mathscr{C}\mathscr{N}(\mathbf{0}, \mathbf{I})}{\sqrt{(1-\alpha)^2 + \alpha^2}}$$

where $\mathscr{C}\mathscr{T}$ is a complex Student-t distribution with degree of freedom 2.1, which has been centered and normalized while $\mathscr{C}\mathscr{N}$ is the classical complex Gaussian distribution.

The set of parameters in this subsection is $p = 3$, $K = 100$ and $\rho = 0.5$.

5.3 Simulations comments

A first comment is that, as expected, the MSE decreases as the estimators make use of more prior informations. On each graphic, the estimator EL1 that does not use any prior on the structure has MSE higher than the MSE of the estimators that make use of the structure informations (EL2 and EL3). These comments are still valid under Gaussian and non-Gaussian assumptions.

Then, we consider the figure 1 which deals with Gaussian distribution. One can see that all estimators have an MSE which decreases as the SNR increases. It is also interesting to notice that the difference between all curves is the same for all SNRs. Moreover, for small SNRs (i.e., for negative SNRs), the MSEs have high values. This can be explained by the fact that the additive noise corrupt the estimation process.

On figure 2, one can notice the improvement introduced by the proposed method in comparison with the classical ML method which degrades for small values of α, i.e. for highly non-Gaussian noise and for a SNR of -17dB.

6. CONCLUSION

This contribution presents an Empirical Likelihood based approach for the identification of multichannel FIR filters in data aided context. The Empirical Likelihood is a flexible semi-parametric estimation method, which can easily integrate in the estimation procedure some prior informations on the structure of the parameter of interest. Moreover, this improved estimation method does not assume any model for the data distribution. The numerical simulation evidence that the proposed method yield reasonable estimates of the channel coefficients even in non-Gaussian additive noise.

7. ACKNOWLEDGEMENT

This work was supported by the French Aerospace Lab through the PEA DTC, the European network of excellence NEWCOM++ and the CNRS/FAPESP program.

REFERENCES

[1] F. Pascal, H. Harari-Kermadec, and P. Larzabal, "The empirical likelihood: an alternative for signal processing estimation," *IEEE Trans.-SP (submitted)*, 2008.

[2] H. Harari-Kermadec and F. Pascal, "On the use of empirical likelihood for non-gaussian clutter covariance matrix estimation," in *Accepted to the IEEE-RADAR 2008*, Roma, June 2008.

[3] J. G. Proakis, *Digital Communications*. McGraw-Hill, Third Ed., New York, 1995.

[4] E. Moulines, P. Duhamel, J. F. Cardoso, and S. Mayrargue, "Subspace methods for the blind identification of multichannel FIR filters," *IEEE Trans. on Signal Processing*, vol. 43, no. 2, pp. 516–525, Feb. 1995.

[5] A. B. Owen, *Empirical Likelihood*. Chapman & Hall/CRC, Boca Raton, 2001.

[6] J. P. Burg, D. G. Luenberger, and D. L. Wenger, "Estimation of structured covariance matrices," *Proc. IEEE*, vol. 70, no. 9, pp. 963–974, September 1982.

[7] D. R. Fuhrmann, "Application of toeplitz covariance estimation to adaptive beamforming and detection," *IEEE Trans.-SP*, vol. 39, no. 10, pp. 2194–2198, October 1991.