A MAXIMUM LIKELIHOOD APPROACH FOR THE PASSIVE IDENTIFICATION OF TIME-VARYING MULTIPATH CHANNELS

Joël Grouffaud, Pascal Larzabal and Henri Clergeot

LESiR - ENS de CACHAN - URA CNRS D1375
61, avenue du Président Wilson - 94235 CACHAN - FRANCE
☎ +33 01 47 40 27 09 - Fax : +33 01 47 40 27 08
joel.grouffaud@lesir.ens-cachan.fr

ABSTRACT

Transmissions through multipath channels suffer from Rayleigh fading and intersymbol interference. This can be overcome by sending a (known) training sequence and identifying the channel (*active* identification). However, in a nonstationary context, the channel model has to be updated by periodically sending the training sequence, thus reducing the transmission rate. We address herein the problem of *blind* identification, which does not require such a sequence and allows a higher transmission rate. In order to track nonstationary channels, we have derived an adaptive (Kalman) algorithm which directly estimates the entire set of characteristic parameters. An original adaptive estimation of the noise model has been proposed for this investigation. Monte-Carlo simulations confirm the expected results and demonstrate the performance.

1. INTRODUCTION

Mobile communications in urban areas are subjected to multipath propagation. This causes Rayleigh fading and intersymbol interference, thereby deteriorating the transmission. Knowledge of the propagation channel is then crucial in order to perform an effective equalization.

The *active* identification consists of sending a known training sequence, whose corresponding response through the channel leads, after a parametric estimation, to the propagation conditions. Propagation channels are almost always time-varying: estimated parameters have to be updated by periodically sending the training sequence, reducing thereby the effective transmission rate.

A two-stage algorithm for the *blind* identification of a stationary multipath channel has been proposed in [1], but can not easily cope with time-varying channels. This paper describes the Conditional Maximum Likelihood (CML) method applied to parametric estimation and provides a Gauss-Newton procedure for implementation. In addition, an original adaptive scheme is proposed for tracking nonstationary channels.

The organization of this paper is as follows. In Section 2, we propose a parametric model of the channel and describe the problem formulation. Section 3 introduces the corresponding log-likelihood and provides an ML estimate of the parameters. In Section 4, a recursive implementation of ML estimation, based on estimation and covariance update, is proposed. A generalization for time-varying parameters is proposed in Section 5 like an adaptive Kalman filtering. Finally, concluding remarks are presented in Section 6.

2. STATEMENT OF THE PROBLEM

2.1 A parametric model of propagation

A radio channel can be depicted by M paths each one characterized by τ_m, θ_m, Δ_m (with $\tau_1 \leq \cdots \leq \tau_M$), which are the group delay, azimuth and elevation angles, respectively, while α_m and ϕ_m are the corresponding attenuation and phase. The emitted signal is supposed to be received by an N-sensor array. The complex, low-pass impulse response between the emitter and the i^{th} sensor, noted $h_i(t)$, is [1, 2]:

$$h_i(t) = \sum_{m=1}^{M} \alpha_m a_i(\theta_m, \Delta_m)\delta(t - \tau_m)e^{j\phi_m} \qquad i = 1 \leq \ldots \leq N, \quad (1)$$

where $a_i(\theta, \Delta)$ is the response of the i^{th} sensor in the bearings (θ, Δ) relative to a reference sensor. It has been assumed in this work that the incoming wavefronts are plane.

In this paper, we assume that the channel is ideal; its response is uniform within the passband and zero elsewhere. The numerical impulse response (finite and causal) is then [1]:

$$\begin{cases} h_i[k] = F_s \sum_{m=1}^{M} \alpha_m a_i(\theta_m) e^{j\phi_m} \operatorname{sinc}(k - F_s \tau_m) & 0 \leq k \leq L \\ h_i[k] = 0 \ \text{for} \ k < 0 \ \text{or} \ k > L \end{cases} . \quad (2)$$

2.2 Representation of spatio-temporal data

The output of the i^{th} sensor at the time k is given by convolving the input signal s[k] and the impulse response, and is further perturbed by an additive noise n_i:

$$x_i[k] = \sum_{l=0}^{L} h_i[l]s[k - l] + n_i[k]. \quad (3)$$

The data can be written under vectorial shape as the set of the outputs of the i^{th} sensor for K successive snapshots:

$$\mathbf{x}_i = \left[x_i[K] \quad \cdots \quad x_i[1]\right]^T, \tag{4}$$

and then gathered into a spatio-temporal vector of length NK:

$$\mathbf{x} = \left[\mathbf{x}_1^T \quad \cdots \quad \mathbf{x}_N^T\right]^T. \tag{5}$$

which may be written as follows:

$$\mathbf{x} = \mathcal{H}\, \mathbf{s}[K.K+L] + \mathbf{n}. \tag{6}$$

$\mathcal{H}$ is a generalized Sylvester matrix of size NK×(K+L), made up of the N impulse responses [3], while **s** is a vector containing signal values. According to (2), $\mathcal{H}$ is actually a function $\mathcal{H}(\Theta)$, where Θ is the vector of unknown spatio-temporal parameters.

2.3 Formulation of the problem

Given only the data collected on the array of sensors, we intend to *identify* the propagation channel (characterized by the set of physical parameters Θ) in a *blind* way, that is without knowing the transmitted signals, whose estimation is furthermore achieved after channel *equalization*. Tracking of the propagation channel (almost always time-varying) has to be considered by the use of an *adaptive* algorithm.

Remark 1

The channel model, defined by (2), is characterized by integer parameters. Some of these are unknown (M and L) and depend on propagation conditions, and may be time-varying. We are assuming in this paper that these parameters are known.

Remark 2

From the shape of the Input/Output relationship (6), it turns out that multiplying each gain $\alpha_m e^{i\phi_m}$ by any complex scalar β amouts to working with the signal $\beta s(.)$. So we shall take in the following simulations $\hat{\alpha}_1 e^{i\hat{\phi}_1} = 1$ (that is $\hat{\alpha}_1 = 1$ and $\hat{\phi}_1 = 0$). Moreover, estimating the M group delays separately is not possible, because no temporal reference is available. Actually, only the differential group delays can be estimated. This can be easily accomplished by imposing $\hat{\tau}_1 = 0$. The set of the spatio-temporal parameters can be written:

$$\Theta = \left[\theta_1 \quad \Delta_1 \;\vdots\; \alpha_2 \quad \phi_2 \quad \theta_2 \quad \Delta_2 \quad \tau_2 \;\vdots\; \cdots \quad \theta_M \quad \Delta_M \quad \tau_M\right]^T$$

3. MAXIMUM LIKELIHOOD ESTIMATION

3.1 Conditional Maximum Likelihood method (CML)

The input s[k] is considered as an unknown, with the parameters to be estimated being the physical parameters of the impulse responses. The only random data element is the additive noise. If many independent observations $\mathbf{x}_p$ (p = 1...P) are available, the covariance matrix takes the form:

$$\hat{\mathbf{R}}_x = \frac{1}{P}\sum_{p=1}^{P} \mathbf{x}_p \mathbf{x}_p^H. \tag{7}$$

Under the assumption of white (both spatially and temporally) gaussian noise, of power σ^2, the maximum likelihood estimation of Θ can easily be derived:

$$\Theta = \arg\max_{\Theta}\left[\mathrm{tr}\left(\Pi_{\mathcal{H}}\hat{\mathbf{R}}_x\right)\right] = \arg\min_{\Theta}\left[\mathrm{tr}\left(\Pi_{\mathcal{H}}^{\perp}\hat{\mathbf{R}}_x\right)\right], \tag{8}$$

where $\Pi_{\mathcal{H}} = \mathcal{H}(\mathcal{H}^H\mathcal{H})^{-1}\mathcal{H}^H$ is the projector on the subspace spanned by the columns of $\mathcal{H}$ ("signal subspace"), and $\Pi_{\mathcal{H}}^{\perp}$ is the projector on the orthogonal subspace ("noise subspace"). This criterion depends no longer on **s**, and the minimization is thus first performed with respect to Θ. Afterwards, the equalization can simply be achieved by using the following relation:

$$\hat{\mathbf{s}} = (\mathcal{H}^H\mathcal{H})^{-1}\mathcal{H}^H\mathbf{x}. \tag{9}$$

Remark 3

It is obvious, from (8), that $\mathcal{H}$ has to be full column rank so that $\mathcal{H}^H\mathcal{H}$ is not singular. This can be strongly connected to the feasability of the blind deconvolution of FIR multichannel and to the coprimness condition for the polynomials $H_n(z)$ $1 \le n \le N$.

3.2 Optimization procedure

This problem can be related to the well-known M.L. search of directions-of-arrival (D.O.A.) in array processing. Optimization of this problem can be conducted by some conventional methods like Wax & Ziskind [4], E.M. [5], and Gauss-Newton. Usually, the final stage of optimization is a Gauss-Newton descent. Once convergence has been achieved, we wish to track variations of the channel. In the next sections, we propose a scheme of blind parametric tracking with an M.L. method.

4. RECURSIVE ML IMPLEMENTATION

4.1 Preliminary: combination of efficient estimates

Let's note Θ_1 and Θ_2 as two independent efficient estimations of the vector Θ; they are assumed to follow normal laws, $\mathcal{N}(\Theta,\Gamma_1)$ and $\mathcal{N}(\Theta,\Gamma_2)$ respectively. For efficient estimation, the log-likelihood is quadratic, and the value Θ_0 of Θ that maximizes $p(X|\Theta_1).p(X|\Theta_2)$ is easily obtained by maximizing the sum of log-likelihoods given by:

$$L(\Theta) = (\Theta-\Theta_1)^H\Gamma_1^{-1}(\Theta-\Theta_1) + (\Theta-\Theta_2)^H\Gamma_2^{-1}(\Theta-\Theta_2). \tag{10}$$

The identification with $(\Theta-\Theta_0)^H\Gamma_0^{-1}(\Theta-\Theta_0)$ leads to:

$$\Theta_0 = (\Gamma_1^{-1}+\Gamma_2^{-1})(\Gamma_1^{-1}\Theta_1 + \Gamma_2^{-1}\Theta_2), \tag{11a}$$

$$\Gamma_0^{-1} = \Gamma_1^{-1} + \Gamma_2^{-1}. \tag{11b}$$

4.2 Combination of one estimate with a new observation

As opposed to the previous case, only one efficient estimate and a new set of observations are available (in fact, a single spatio-temporal observation is sufficient). Indeed, the new observations are unable to provide alone an efficient estimation for the parameters. We assume that Θ is normally distributed with the density being, at time t, $\mathcal{N}(\hat{\Theta}_t, \Gamma_t)$. Furthermore, the log-likelihood at time t can be written according to its second-order expansion (quadratic approximation) using the gradient and the hessian of the log-likelihood:

$$\mathcal{L}(X|\Theta) = \nabla_t^H(\Theta - \hat{\Theta}_t) + \frac{1}{2}(\Theta - \hat{\Theta}_t)^H \mathbf{H}_t(\Theta - \hat{\Theta}_t). \quad (12)$$

It is assumed that the resulting log-likelihood function (after integration of the new measurements) is sufficiently sharply-peaked around $\hat{\Theta}_t$, so that the second-order approximation remains valid over the interval $\left[\hat{\Theta}_t, \hat{\Theta}_{t+1}\right]$. The resulting log-likelihood (up to a constant term) is obtained from Bayes integration formula:

$$\mathcal{L}(\Theta|X) = \mathcal{L}(X|\Theta) + \frac{1}{2}(\Theta - \hat{\Theta}_t)^H \Gamma_t^{-1}(\Theta - \hat{\Theta}_t). \quad (13)$$

The identification with the *a posteriori* Gaussian distribution $\mathcal{N}(\hat{\Theta}_{t+1}, \Gamma_{t+1})$ leads to the maximum a posteriori (MAP) estimate:

$$\begin{cases} \hat{\Theta}_{t+1} = \hat{\Theta}_t - \Gamma_t . \nabla_t \\ \Gamma_{t+1}^{-1} = \Gamma_t^{-1} + \mathbf{H}_t \end{cases} \quad (14)$$

Figure 1 provides an overview of the recursive implementation of the maximum likelihood method for the update of estimation and covariance. Both gradient and Hessian are calculated from the new data alone , and are combined with a priori estimate (at time t) to give the a posteriori estimate (at time t+1), according to (14).

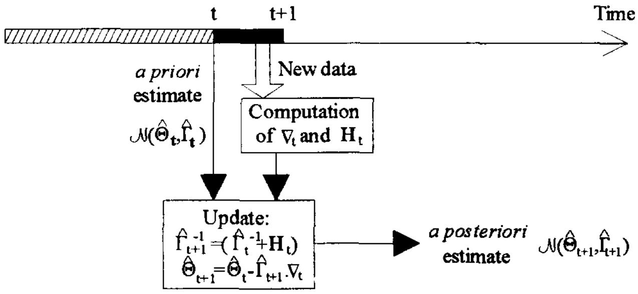

Figure 1: Maximum Likelihood and recursive implementation.

4.3 Simulations

Simulations were carried out to demonstrate the applicability and the performance of the proposed algorithm. All of the following simulations have been conducted for a 5-element linear array with equispaced sensors. A spatio-temporal observation has been built from K=20 samples (it is thus a 100-dimensional vector). Five snapshots were used to compute the covariance matrix (7), and the gradient and Hessian. The input signal type is FSK, and the received data are sampled at four times the baud rate. The channel is represented by two paths, whose main parameters are given in the following table:

m	α_m	ϕ_m	θ_m	Δ_m	$\tau_m F_s$
1	1	0	90 °	30 °	0.2
2	0.7	20 °	90 °	55 °	2.9

The order of the impulse response is assumed to be known in this study (see remark 1): L=4. Root-Mean-Square Error (RMSE) is used to measure the accuracy of parameters estimates:

$$\text{RMSE} = \sqrt{\sum_{i=1}^{\text{length}(\Theta)} \left(\frac{\Theta_i - \hat{\Theta}_i}{\Theta_i} \right)^2} . \quad (15)$$

We have studied the convergence of this algorithm by plotting on figure 2 RMSE versus iteration number for many values of the Signal to Noise Ratio (SNR).

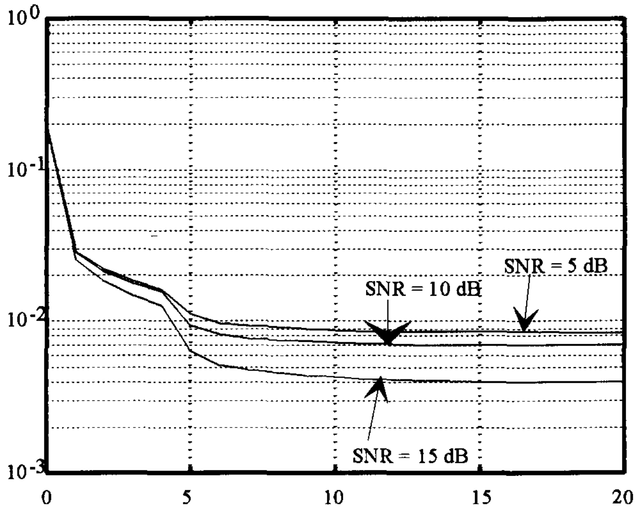

Figure 2: Convergence of the recursive algorithm.

5. AN ADAPTIVE ALGORITHM FOR TIME-VARYING CHANNELS

5.1 Dynamic model

For time-varying channels, a dynamic model for the evolution of parameters is assumed:

$$\Theta_{t+1} = \Theta_t + \mathbf{u}_t + \varepsilon_t. \quad (16)$$

$\mathbf{u}_t$ accounts for the long-term average motion, and ε_t for fast decorrelated position increments (random walk). In fact, $\mathbf{u}_t$ can be seen as a command while ε_t may be the process noise, zeromean, uncorrelated with Θ and $\mathbf{u}$, and with $\mathbf{R}_\varepsilon = E\left[\varepsilon_t \varepsilon_t^H\right]$.

5.2 Algorithm derivation

With the "command" $\mathbf{u}_t$ being assumed to be known, the best estimate of Θ at time t+1, without any new measurement, is:

$$\hat{\Theta}_{t+1|t} = \hat{\Theta}_{t|t} + \mathbf{u}_t. \quad (17)$$

The error of estimation is ε_t, and the computation of its covariance matrix, associated with equation (16), leads to:

$$\hat{\Gamma}_{t+1|t} = \hat{\Gamma}_{t|t} + \mathbf{R}_\varepsilon. \quad (18)$$

Assuming that the resulting log-likelihood is sufficiently sharply-peaked to be well approximated by its second-order expansion on the interval $\left[\hat{\Theta}_{t+1|t}, \hat{\Theta}_{t+1|t+1}\right]$, and considering the set of equations (14), (17) and (18), it is possible to propose the prediction-estimation filter:

$$\begin{cases} \hat{\Theta}_{t+1|t} = \hat{\Theta}_{t|t} + \mathbf{u}_t \\ \text{Computation of } \mathbf{H}(\hat{\Theta}_{t+1|t}) \text{ and } \nabla(\hat{\Theta}_{t+1|t}) \\ \hat{\Gamma}_{t+1|t} = \hat{\Gamma}_{t|t} + \mathbf{R}_\varepsilon \\ \hat{\Gamma}_{t+1|t+1}^{-1} = \hat{\Gamma}_{t+1|t}^{-1} + \mathbf{H}(\hat{\Theta}_{t+1|t}) \\ \hat{\Theta}_{t+1|t+1} = \hat{\Theta}_{t+1|t} - \hat{\Gamma}_{t+1|t+1} \cdot \nabla(\hat{\Theta}_{t+1|t}) \end{cases} \quad (19)$$

However, the previous filter does not remain efficient in the most general situation, because neither $\mathbf{u}_t$ nor $\mathbf{R}_\varepsilon$ are known, and they are generally time-varying.

5.3 Recursive noise covariance computation

The simplest way to estimate $\mathbf{u}_t$, by using the filtered positions, is:

$$\hat{\mathbf{u}}_t = \hat{\Theta}_{t|t} - \hat{\Theta}_{t-1|t-1}. \quad (20)$$

However, this kind of estimation is very sensitive to statistical fluctuations and may degrade performance. The proposed solution consists of estimating $\mathbf{u}$ through a finite time averaging (smoothed estimation) according to:

$$\hat{\mathbf{u}}_t = (1-\mu_1)\hat{\mathbf{u}}_{t-1} + \mu_1(\hat{\Theta}_{t|t} - \hat{\Theta}_{t-1|t-1}). \quad (21)$$

The covariance matrix of the random walk is recursively estimated in the same way, by using the covariance of the gradient [6]:

$$\mathbf{R}_t = \mathbf{H}_t^{-1}\operatorname{cov}(\nabla_t)\mathbf{H}_t^{-1} - \hat{\Gamma}_{t|t} - \mathbf{H}_t^{-1}. \quad (22)$$

However, the covariance of the gradient is untractable, and we shall use a smoothed estimation:

$$\hat{\mathbf{R}}_t = (1-\mu_2)\hat{\mathbf{R}}_{t-1} + \mu_2(\mathbf{H}_t^{-1}\nabla_t\nabla_t^H\mathbf{H}_t^{-1} - \hat{\Gamma}_{t|t} - \mathbf{H}_t^{-1}). \quad (23)$$

The whole Kalman filter for the adaptive estimation of Θ is merely derived from equations (19) by replacing $\mathbf{u}_t$ and $\mathbf{R}_\varepsilon$ by $\hat{\mathbf{u}}_t$ and $\hat{\mathbf{R}}_t$ respectively, as estimated on line according to (21) and (23).

5.4 Simulations

This second simulations set illustrates the behaviour of the adaptive algorithm when the channel is time-varying. We consider a slowly time-varying channel, for which the evolution is determistic, according to a sine wave. The context of simulations is the same as in 4.3 for fixed parameters. SNR is fixed at 12 dB. The initial conditions are near from the true values of parameters, that is we assume that convergence as almost been achieved. Figure 3 shows the tracking of such a channel.

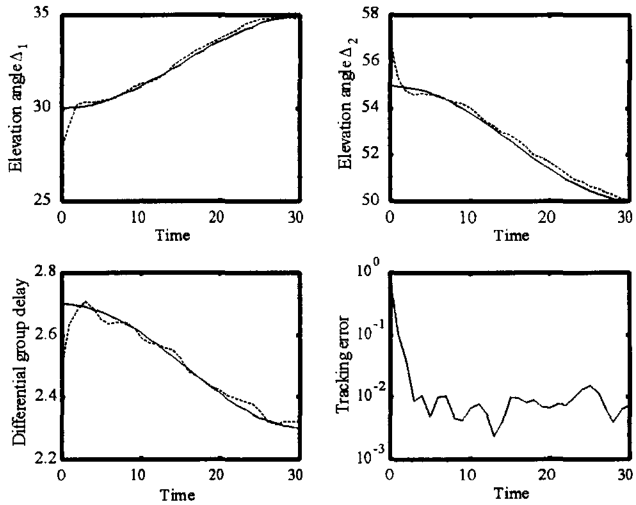

Figure 3 : Tracking of a time-varying channel.
(solid line : true trajectory - dotted line : estimated trajectory)

6. CONCLUSION

We have proposed an algorithm for the blind parametric estimation of time-varying multipath channels. Statistical performance and tracking properties have been studied by numerical simulations. Future work involves a theoretical analysis of the convergence properties and a comparative study with other blind tracking algorithms.

REFERENCES

[1] J. Grouffaud, A. Ferréol, P. Larzabal, H. Clergeot, "Passive Identification of Multipath Channels", *Proceedings of EUSIPCO*, September 1996.

[2] J. Proakis, *Digital Communications*, McGraw-Hill, New-York, 1989.

[3] E. Moulines, P. Duhamel, J.F. Cardoso and S. Mayrargue, "Subspace Methods for the Blind Identification of Multichannel FIR Filters", *IEEE Transactions on Signal Processing*, Vol. 43, No. 2, February 1995.

[4] I. Ziskind, M. Wax, "Maximum Likelihood Localization of Multiple Sources by Alternating Projection", *IEEE Transactions on ASSP*, Vol. 38, No. 10, October 1988.

[5] M. Miller, D. Fuhrman, "Maximum Likelihood Narrow-Band Direction Finding and the E.M. Algorithm", *IEEE Transactions on ASSP*, Vol. 38, No. 9, September 1990.

[6] J. Grouffaud, A. Ferréol, P. Larzabal, H. Clergeot, "Adaptive Maximum Likelihood Algorithms for the Blind Tracking of Time-Varying Multipath Channels", submitted to International Journal on Adaptive Control and Signal Processing.