



Is training of adaptive equalizers still useful ?

Joël Labat, Odile Macchi, Christophe Laot, N. Le Squin

► **To cite this version:**

Joël Labat, Odile Macchi, Christophe Laot, N. Le Squin. Is training of adaptive equalizers still useful ?. Globecom'96, Nov 1996, Londres, United Kingdom. pp.968 - 972. hal-02158754

HAL Id: hal-02158754

<https://hal.archives-ouvertes.fr/hal-02158754>

Submitted on 18 Jun 2019

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Is Training of Adaptive Equalizers still Useful ?

J. Labat *, O. Macchi **, C. Laot * and N. Le Squin*

* FRANCE TELECOM / ENST B, BP 832, 29285 BREST Cédex FRANCE

** LSS/CNRS/ESE, Plateau de Moulon, 91192 Gif sur Yvette Cédex FRANCE

Tel : 98 00 13 65 Fax : 98 00 10 12 E.mail : joel.labat@enst-bretagne.fr

Abstract - We present a novel unsupervised adaptive equalizer. It has the same computational complexity, convergence speed and steady-state MSE as a trained LMS adaptive DFE, but it is not subject to error propagation. Therefore it can equalize even severe and/or quickly varying channels. This follows from the very structure of the equalizer, which allows a completely reversible transition between (i) a linear structure in the starting mode : the decoupled cascade of a recursive adaptive predictor and a transversal phase equalizer and (ii) a classical DFE in the tracking mode. The equalizer behaviour is fully satisfactory during hours of real underwater communications. It reaches the standard of trained equalizers. Hence the question in the title.

I. INTRODUCTION

Let $d(n)$ be a zero-mean, unit power, independent sequence of discrete data to be transmitted through a discrete channel \mathcal{F} with transfer function $F(z)$. In this paper, $d(n)$ and \mathcal{F} can be complex-valued to account for QAM transmission. As a consequence, the resulting signal $s(n)$ needs to be equalized. More precisely three effects have to be corrected, namely amplitude distortion, phase distortion and gain. Hence, the equalizer can be split into three cascaded devices, as already suggested in [1].

Amplitude equalization (AE). This is realized by a whitening filter, or innovator. For reasons which will become apparent later, here we select a purely recursive whitener with transfer function (TF) $1/(1+A(z))$, expected to be causal and stable.

Phase equalization (PE). This is performed by an all-pass filter, which we choose as a FIR filter. Denote $P(z)$ its transfer function.

Complex gain. It is carried out by (i) an automatic gain control (GC) which adjusts the output power to the unitary power level of $d(n)$ and (ii) a digital phase locked loop (PLL) in order to cope with the residual phase error that remains after the demodulation.

Clearly, with fixed devices, the order for implementing the (AE), (PE), (GC) and (PLL) is irrelevant. However with adaptive filters, and with a non stationary channel \mathcal{F} , the order will turn out to be critical.

The purpose of this contribution is to present a novel and very efficient method to adaptively control these three devices in an unsupervised (blind) fashion : no training sequence $d(n)$ is required to optimize the three devices.

A first approach of this kind has been presented in [1], [2]. However, when a zero $z_{0,j}$ of $F(z)$ outside the unit circle is close to it, the system stability is poor due to the choice of a recursive structure for the (PE). Here we improve the results of [1] by choosing a FIR (PE). Moreover we grant the system a soft structure in that the places of (AE) and (PE) can be switched at any time, based on the result of some simple performance test. In severe situations originating, e.g., in sudden changes of the channel \mathcal{F} , the (AE) comes before the (PE) and the two updates are decoupled, both equalizers being controlled in an unsupervised mode. When the channel equalization is sufficiently improved, the (PE) comes first, the (AE) is switched in a decision directed mode, both equalizers being jointly controlled in the classical decision directed mode. With this structure the system is equivalent to a classical DFE. The advantage is that the change of structure is perfectly reversible, which makes the system able to track severe non stationary channels without being ever stopped by the well-known phenomenon of error propagation which destroys the benefit of DFE when the channel suddenly changes. The robustness gained in this way costs no additional computational complexity : the updating needs essentially the same amount of multiplications as a classical adaptive DFE and no division is needed.

In this way, the novel adaptive equalizer appears as the most efficient one for severely non stationary channels.

II STARTING PERIOD : (AE) BEFORE (PE)

II-1 The structure. This section presents a structure which can be self-adaptively controlled even though it starts with a "closed eye". Clearly the (PE) only acts on the phase of the channel TF. Therefore, irrespective of its place, the input signal of the (AE) has the same spectrum.

The role of the (AE) is to equalize this spectrum : its output must be white (flat spectrum). Equivalently, the output sequence must be uncorrelated. This is obtained in a self-adaptive manner by means of prediction. Therefore the (AE) can be easily optimized, independently of its place, with a control based on local (input or output) signals.

A similar property does not hold for the (PE). In particular if it comes first, the (PE) optimality is not characterised by properties of its own output (like independence of the sequence or acquisition of a given distribution), but by properties of the final signal at the (AE) output. Therefore adaptations of the (PE) and (AE) are coupled. As a result if the (AE) is poor, the (PE) will suffer from a pump effect.

This is why to start convergence, it is better to place the (AE) first. Then the (PE) will enjoy a local control, solely based on its own input and output. Hence the system starts with the structure depicted in Fig. 1 where the notations are given. The signals $s(n)$, $t(n)$, $u(n)$, $v(n)$, $w(n)$ and $\hat{d}(n)$ are complex valued.

The (GC) can be implemented at any place. We put it first to decrease the dynamical ranges required in the other two devices. As shown in the figure, we choose a purely recursive filter for the (AE). This choice will be legitimated in section III : it will allow a great performance improvement in the tracking period. The (PE) section is taken as an FIR filter. In the case of complex signals, it is followed by the (PLL) which copes with the residual phase error affecting the signal $v(n)$ before the threshold decision device (DD).

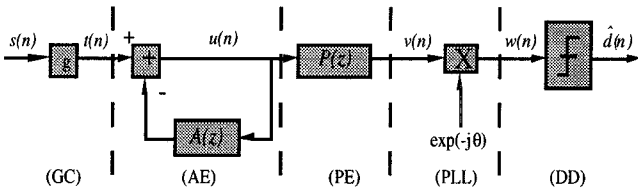


Figure 1 Linear starting structure (convergence period)

The (AE) time equation is

$$u(n) = t(n) - \mathbf{A}^T \mathbf{U}_N(n-1) \quad (1)$$

where the vector $\mathbf{A} = [a_1, \dots, a_{N-1}, a_N]^T$ characterizes the (AE) and

$$\mathbf{U}_N(n-1) = (u(n-1), \dots, u(n-N))^T \quad (2)$$

is made of N past samples of $u(n)$ in such a way that $\mathbf{A}^T \mathbf{U}_N(n-1)$ is a prediction of $t(n)$ and $u(n)$ is the prediction error (or innovation). Optimization of the (AE) will simply consist in choosing \mathbf{A} such that the output $u(n)$ is an uncorrelated sequence.

The (PE) time equation is

$$v(n) = \mathbf{P}^T \mathbf{U}_{L+1}(n) \quad (3)$$

where the complex vector $\mathbf{P} = [p_0, p_1, \dots, p_L]^T$ characterizes the (PE) and

$$\mathbf{U}_{L+1}(n) = (u(n), u(n-1), \dots, u(n-L))^T \quad (4)$$

is made of the present sample $u(n)$ plus L past samples. Optimization of the (PE) will consist in choosing \mathbf{P} such that the output $v(n)$ is, e. g., an independent sequence. Other criteria are possible (see below).

The (GC) and (PLL) equations are respectively

$$t(n) = gs(n) \quad , \quad w(n) = v(n) \exp(-j\theta) \quad (5)$$

where g is a positive value and θ is real.

II-2 The adaptation. The (GC). It is easy to design an adaptive (GC) based on the unit power property of the data $d(n)$, see, e. g., [3] ch. 1.

$$G(n) = G(n-1) + \mu_G [1 - |u(n)|^2] \quad , \quad g(n) = \sqrt{|G(n)|} \quad (6)$$

where μ_G is the positive step size. This algorithm ensures that the input power of the (PE) is unitary.

The (AE). It is well-known that decorrelation of the prediction error $u(n)$ is equivalent to minimization of the prediction error power. Hence \mathbf{A} is chosen to minimize $E\{|u(n)|^2\}$. This can be done via a stochastic gradient algorithm which updates the complex parameter \mathbf{A} at time n using increment

$$\Delta \mathbf{A} = -(\mu_A / 2) \nabla_{\mathbf{A}} |u(n)|^2 \quad , \quad \mu_A > 0 \quad (7)$$

For a recursive filter this quantity cannot be exactly computed but efficient approximate algorithms are available [3], ch. 14. For example, the stochastic gradient LMS leads to

$$\mathbf{A}(n) = \mathbf{A}(n-1) + \mu_A u(n) \mathbf{U}_N^*(n-1) \quad (8)$$

$$u(n) = t(n) - \mathbf{A}^T(n-1) \mathbf{U}_N(n-1) \quad (9)$$

where the superscript $*$ stands for complex conjugate. In other words, the adaptive parameter \mathbf{A} is in state $\mathbf{A}(n-1)$ at time n , while (2) remains valid. The same notation is used for all quantities.

The (PE). The optimization criteria to update \mathbf{P} depend on the a priori knowledge about $d(n)$.

(i) $d(n)$ is discrete, with known levels. The usual criterion is the decision directed mean square error (DDMSE)

$$J_{DD}(\mathbf{P}) = E\{|w(n) - \hat{d}(n)|^2\} \quad (10)$$

Basically this criterion assumes that the overall system works well enough for the eye to be open, in such a way

that $\hat{d}(n) = d(n - \delta)$ in (10) where δ is an irrelevant delay. So the (PE) is optimized according to the classical supervised MSE criterion. The stochastic gradient algorithm associated to the criterion (10) easily follows from (3) and (5). It takes the form

$$\mathbf{P}(n) = \mathbf{P}(n-1) + \mu_P e(n) \mathbf{U}_{L+1}^*(n) \quad (11)$$

$$e(n) = e_{DD}(n) = \hat{d}(n) \exp(j\theta(n-1)) - v(n) \quad (12)$$

where μ_P is positive. The theoretical study of the criterion (10) is very awkward because the decisions $\hat{d}(n)$ can be wrong when \mathbf{A} and \mathbf{P} lie in certain regions. Therefore we use the second kind of knowledge.

(ii) the sequence $d(n)$ is independent. The usual criterion is to minimize the fourth moment, subject to a (unitary) power constraint at the output of the (PE), corresponding, e.g., to the Godard criterion [4]

$$J_G(\mathbf{P}) = \mathbf{E}\{|v(n)|^4\} - 2\mathbf{E}\{|v(n)|^2\} \quad (13)$$

which intends to minimizing the fluctuations of $|v(n)|^2 - 1$. The corresponding stochastic gradient algorithm takes the same form (11) but the error is

$$e(n) = e_G(n) = v(n) (1 - |v(n)|^2) \quad (14)$$

Note that in [5], Shalvi and Weinstein minimize $\mathbf{E}\{|v(n)|^4\}$ but the constraint $\mathbf{E}\{|v(n)|^2\} = 1$ is ensured by the all-pass form of their (PE) filter. The associated error $e_{SW}(n) = -v(n)|v(n)|^2$ is not used in our paper where the constraint that (PE) is an all-pass filter is relaxed. The theoretical study of the criterion (13) is still in progress in the open literature. It appears that, provided L is large enough, J_G has only global minima. So the corresponding algorithm (11) and (14) is efficient to recover the true data $d(n)$, up to an arbitrary phase θ and an unknown delay δ . So in steady-state

$$v(n) = d(n - \delta) \exp(j\theta) \quad (15)$$

The (PLL). The phase error correction can be done using the DDMSE criterion. Moreover, the signal error can be filtered via a recursive digital filter in order to correct a possible frequency offset. So the update equations are

$$\varepsilon(n) = \text{Im}\{w(n) [\hat{d}(n) - w(n)]^*\} \quad (16)$$

$$q(n) = G_1 \varepsilon(n) + G_2 \sum_{k=1}^n \varepsilon(k),$$

$$\theta(n) = \theta(n-1) + \mu_\theta q(n) \quad (17)$$

where μ_θ , G_1 and G_2 are real positive parameters to be properly chosen according to the application. This kind of update is clearly decision directed. However, we never encountered any problems related to our simulations when directly starting in this way.

II-3 The switching rule. During this starting period, in order to get some performance measure, the DDMSE is estimated according to (λ denoting the forgetting factor) :

$$J_{DD}(n) = \lambda J_{DD}(n-1) + (1-\lambda) \left| \hat{d}(n) - w(n) \right|^2 \quad (18)$$

When this estimate decreases below a given performance threshold J_0 , the (PE) and the (AE) are permuted according to the Fig. 2 (switch in position b).

III TRACKING PERIOD : (PE) BEFORE (AE)

III-1 The structure During the tracking mode, the equalizer is described by Fig. 2. When $J_{DD}(n)$ decreases below a properly chosen threshold $J_1 \leq J_0$, the recursive part of the (AE) is gradually fed with the past decisions (switch in position a). As a result, our system becomes the classical DFE. In the sequel, we shall assume that $J_1 = J_0$.

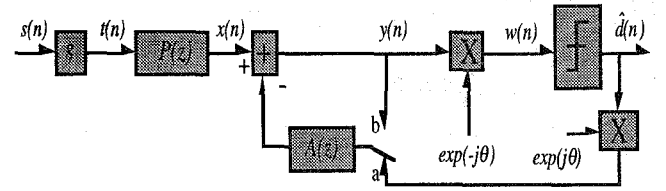


Figure 2 Structure of the equalizer (tracking period)

Obviously this removes a large part of the noise and of the remaining interference still present in the signal $w(n)$ and improves the final error rate. In this configuration, $y(n)$ is replaced by $\hat{d}(n) \exp(j\theta(n-1))$. As a consequence, our system allows soft transitions between a recursive linear self-learning equalizer and a classical DFE using the DDMSE criterion for updating both \mathbf{A} and \mathbf{P} . Clearly this switch between the starting and tracking periods is reversible. This is a very important and attractive feature, specially for the severe non stationary channels. Besides, we can hold the gain $g(n)$ at a fixed value g . Then the (PE) will ensure the (GC) function.

The time equations used during the tracking period are the following

$$t(n) = gs(n) \quad (19)$$

$$w(n) = \left(\mathbf{P}^T(n-1) \mathbf{T}(n) - \mathbf{A}^T(n-1) \mathbf{D}(n) \right) e^{-j\theta(n-1)} \quad (20)$$

where $\mathbf{T}(n) = [t(n), \dots, t(n-L)]^T$ and

$$\mathbf{D}(n) = [\hat{d}(n-1)e^{j\theta(n-2)}, \dots, \hat{d}(n-N)e^{j\theta(n-N-1)}]^T$$

III-2 The adaptation. The simplest updating algorithm is the classical DDLMS one which updates \mathbf{A} and \mathbf{P} according to the stochastic gradient of the criterion (10).

$$\mathbf{P}(n) = \mathbf{P}(n-1) + \mu_B [\hat{d}(n) e^{j\theta(n-1)} - y(n)] \mathbf{T}^*(n) \quad (21)$$

$$\mathbf{A}(n) = \mathbf{A}(n-1) - \mu_C [\hat{d}(n) e^{j\theta(n-1)} - y(n)] \mathbf{D}^*(n) \quad (22)$$

The (PLL) adaptation remains unchanged while the (GC) is now performed in the (AE). By the way, the (PLL) place can be changed, e.g., located between the (PE) and the (AE). This is straightforward.

III-3 Switching rule. In order to go back and forth from one structure to the other, $J_{DD}(n)$ must be permanently evaluated. The switch back to the starting mode occurs when $J_{DD}(n)$ oversteps J_0 .

IV COMPUTER SIMULATIONS

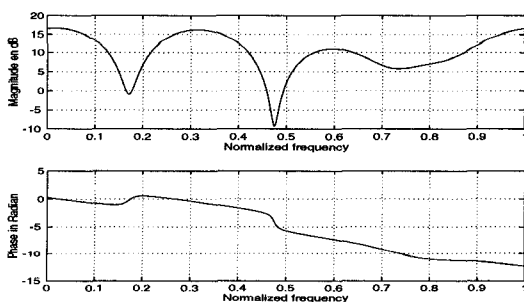
Results have been obtained via Monte Carlo simulations using 200 different runs. We have selected two severe channels. Channel 1 and 2 have been proposed in [6] and in [2] respectively. Their impulse responses are

$$\text{Channel 1 } \mathbf{f}_1 = [2-0.4j, 1.5+1.8j, 1, 1.2-1.3j, 0.8+1.6j]$$

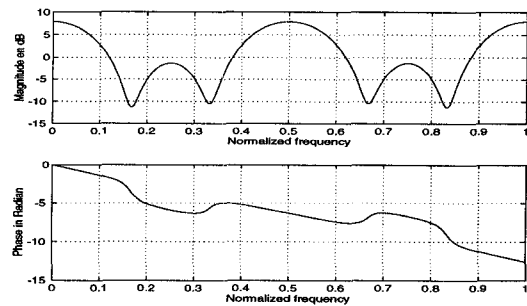
$$\text{Channel 2 } \mathbf{f}_2 = [0.8264, -0.1653, 0.8512, 0.1636, 0.81]$$

Figure 3.a (resp. 3b) illustrates the amplitude and phase response of channel 1 (resp. 2) while figure 4.a (resp. 4.b) depicts the location of the TF zeros of channel 1 (resp. 2).

Note that both channels exhibit deep fading frequencies and severely non linear phase distortion, as a result of the zeros of $F(z)$ which are outside the unit circle and close to it. For such severe channels most self-adaptive (blind) equalizers proposed in the literature do not work at all.



3a Channel 1



3b Channel 2

Figure 3 Amplitude and phase responses

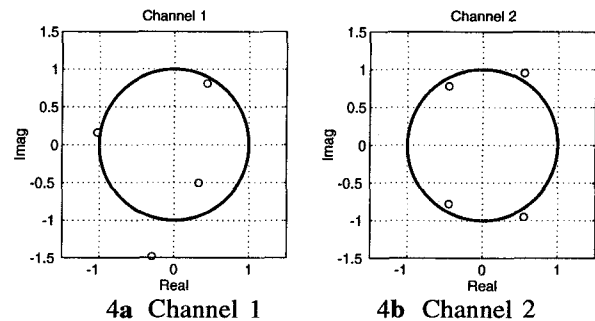


Figure 4 Zeros of the transfer function

Five adaptive equalizers have been tested. Three of them are transversal linear equalizers (LE) with 31 taps, started with a centered reference tap. They are the Godard equalizer [4], the Duhamel and Hilal one [7] and the usual supervised (LE). The two recursive equalizers are the classical supervised (non linear) DFE and the new system of this paper. Both recursive equalizers have 5 and 20 taps in their recursive and transversal parts, respectively. For the 16-QAM scheme, the Duhamel and the Godard unsupervised algorithms are totally inefficient for these two severe channels. So our system is only compared to the supervised (LE) and (DFE) equalizers. The forgetting factor to compute $J_{DD}(n)$ is $\lambda=0.99$, starting with $J_{DD}(0) = 1$, while the switching threshold is 0.2 (-7dB) for the 4-QAM and 0.063 (-12 dB) for the 16-QAM. It means that our system starts with the appropriate structure : (AE) comes first. The equalizer input SNR for the 4-QAM and 16-QAM schemes are 15 dB and 25 dB respectively.

In the following, the impulse responses are normalized to unity. The additive noise is zero mean, white and gaussian, whereas the channel phase shift standing for the demodulation phase error is $\phi(n) = 2\pi n\Delta FT + \Phi$, T denoting the symbol duration.

In the 200 trials, the phase error Φ is uniformly distributed in $[0, 2\pi]$. Moreover the normalized frequency offset ΔFT is uniformly distributed in $[0, 1e-3]$ for 4-QAM and in $[0, 1e-4]$ for 16-QAM. Then the MSE is estimated in an RLS sense (cf 18) for every experiment and averaged over the 200 trials.

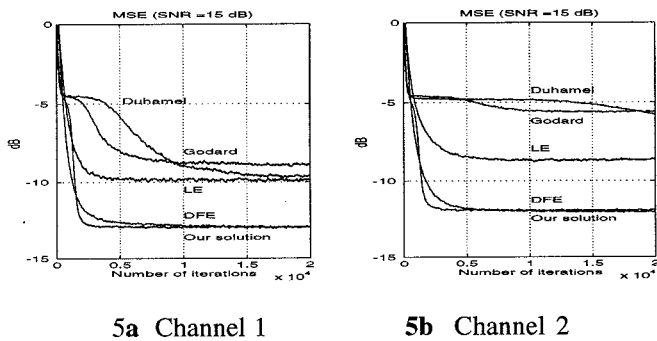


Figure 5 Mean Square Error for 4-QAM

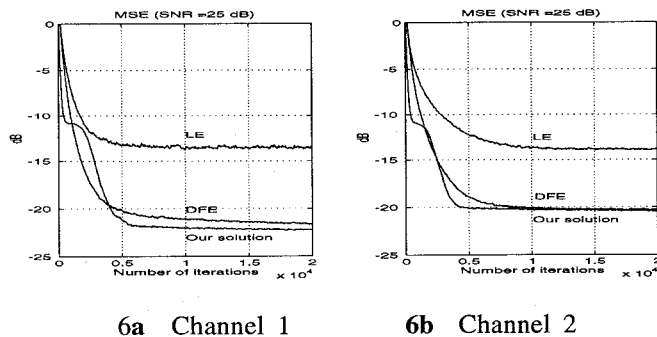


Figure 6 Mean square error for 16-QAM

The striking important conclusion drawn from the simulations of Fig. 5 and 6 is that our new unsupervised system achieves essentially the same performance as the best supervised equalizer which is the DFE. We can even obtain a slightly better achievement !

Of course the classical transversal equalizers are totally outperformed (3 dB to 9 dB of output MSE gain) as a result of the channel severity. Convergence of the new equalizer respectively requires less than 2000 and 5000 iterations, in the 4-QAM and 16-QAM cases. It is very fast.

V CONCLUSION

From the structural point of view, the new equalizer is based on some ideas previously developed by Macchi et al. [1], [2]. The specific new idea proposed in this paper is to permute in a softly reversible way, the two main parts of the equalizer, namely the recursive (AE) and the transversal (PE), according to some performance test, e. g., the DDMSE [8]. In the starting periods the system is linear and unsupervised (blind). In the tracking periods it is a classical adaptive DFE. Hence it is particularly well suited for severe channels which exhibit sudden or fast time variations : whenever this is possible, it takes advantage of the decision directed mode but it never suffers from error propagation and is always able to work in a self-learning mode.

It appears to be the first time that an adaptive DFE is successfully employed in a completely unsupervised manner : no loss of convergence speed is observed

compared to the supervised mode and the steady state MSE is clearly the same. Moreover the computational complexity is not increased (compare equations (8) and (11) to (22) and (21) respectively).

This equalizer has been successfully tested on real underwater communication system with 4-QAM modulation at 6 kbit/s. The shallow underwater acoustic channel is a very severe one suffering from both multipath effects and Doppler shifts. At this rate, the impulse response may spread over 40 symbol durations. Despite of all these channel impairments the new system never got lost during two hours.

This success is enough convincing in itself to raise the question of our title : is the training period still worthy for LMS adaptive equalizers ?

REFERENCES

- [1] C. A. Faria Da Rocha, O. Macchi, J. M. T. Romano, "An adaptive nonlinear IIR filter for self-learning equalization", ITC 94, Rio-de-Janeiro, Brazil, pp.6-10.
- [2] O. Macchi, C. A. Faria Da Rocha, J. M. T. Romano, "Egalisation adaptative autodidacte par rétroprédiction et prédiction", 14th GRETSI Symposium, Juan-Les-Pins, pp 491-494, 13-16 Septembre 1993.
- [3] O. Macchi "Adaptive Processing. The LMS Approach with Applications in Transmission" Wiley, 1995.
- [4] D. N. Godard, "Self-recovering equalization and carrier tracking in two dimensional data communication system", IEEE Trans. on Com., vol. 28, 1980, pp. 1867-1875.
- [5] O. Shalvi, E. Weinstein, "New criteria for blind deconvolution of nonminimum phase systems", IEEE Trans. on IT, vol. 36, No 2, March 1990, pp. 312-321.
- [6] B. Porat, B. Friedlander, "Blind equalization of digital communication channels using high-order moments", IEEE Trans. on SP, vol. 39, 1991, pp. 522-526.
- [7] K. Hilal, P. Duhamel "A blind equalizer allowing soft transition between the CMA and the DD algorithm for PSK modulated signals" ICC, pp 1144-1148, Geneva, May 1993.
- [8] J. Labat, C. Laot, O. Macchi, "Dispositif d'égalisation adaptatif pour systèmes de communications numériques", French Patent 9510832 (15 september 95).