A PDF Estimation-Based Blind Criterion for Adaptive Equalization

Charles C. Cavalcante[1], F. Rodrigo P. Cavalcanti[2] and João Cesar M. Mota[2]

[1] UNICAMP, Campinas-SP, Brazil

[2] UFC, Fortaleza-CE, Brazil

Abstract — **A blind criterion for adaptive equalization based on probability density function (pdf) estimation is proposed. The criterion measures the divergence between the pdf of an ideally equalized signal and the one from a parametric model resulting in a cost function that is a sort of entropy minimization of the equalizer output signal. It is also shown a link between the constant modulus (CM) criterion and the proposed one under certain circumstances. Some convergence properties are studied and the performance of the proposal is evaluated through simulations faced to a classical blind criterion.**

I. Introduction

In the design of receivers in digital communication systems blind receivers are being more and more used. Since the work of Sato [1] and lately the one of Godard [2] the study on blind criteria has increased a lot.

The most part of proposed blind criteria take in account the Bussgang methods, it means, a memoryless nonlinear function is applied to the equalizer output to produce a signal with the same statistical properties of the transmitted ones in order to compare it with the equalizer output [3].

Some recent works have proposed information-theoretic based approaches for blind equalization criteria [4], [5]. In those papers, the nonparametric estimation of the probability density function (pdf) of the signal on the equalizer output is required for the cost function construction, besides, a nonlinear structure for the equalizer, such as neural networks, is used [4]. Another characteristic of information-theoretical approaches is their use in blind source separation, which has a strong link with blind equalization (deconvolution) [6]

The present proposal is information-theoretical based with parametric pdf estimation. The pdf of the ideally (perfectly) equalized signal is *a priori* determined based on feasible assumptions, making possible to use a parametric model that fits the system order. Hence, the resulting criterion is based on the entropy minimization of the equalizer output signal.

An stochastic gradient type algorithm is derived for the updating of the filter equalizer. A brief analyze of the stochastic version of the constant modulus (CM) criterion is done in order to show a link between that criterion and the proposal.

C.C. Cavalcante is with the Digital Signal Processing for Communications (DSPCom) Lab., Dep. of Communications, Faculty of Electrical and Computer Engineering, Campinas State University (UNICAMP), Phone: +55 19 3788 3702, Fax: +55 19 3289 1395, C.P. 6101, CEP: 13083-970, Campinas-SP, Brazil. (email: `charles@decom.fee.unicamp.br`)

F.R.P. Cavalcanti and J.C.M. Mota are with the GTEL Lab., Dep. of Teleinformatics Engineering, Federal University of Ceará (UFC), C.P. 6007, CEP: 60755-460, Fortaleza-CE, Brazil. (emails: {`rod,mota`}`@gtel.ufc.br`)

This work was supported by the Ericsson Research Brazilian Branch under the ERBB/UFC Technical Cooperation Contract.

The used notation is given as follows. The discrete transmitted sequence $\mathbf{a}(n) = \left[\begin{array}{ccc} a(n) & \cdots & a(n-N-M+1) \end{array} \right]^T$ is assumed independent and identically distributed (i.i.d) and the symbols $a(n) \in \mathcal{A}$ which has cardinality S. The channel is represented by a FIR filter $\mathbf{h} = \left[\begin{array}{ccc} h_0 & \cdots & h_{N-1} \end{array} \right]^T$. Additive noise is white, Gaussian, uncorrelated from the transmitted sequence, its variance σ_n^2 is given according the signal-to-noise ratio (SNR) and will be denoted by $\mathbf{v}(n) = \left[\begin{array}{ccc} v(n) & \cdots & v(n-M+1) \end{array} \right]^T$. The equalizer, which has finite impulse response (FIR) denoted by $\mathbf{w}(n) = \left[\begin{array}{ccc} w(n) & \cdots & w(n-M+1) \end{array} \right]^T$, is feed by the channels outputs $x(n) = \bar{x}(n) + v(n)$ where $\bar{x}(n) = \sum_{i=0}^{N-1} h_i a(n-i)$ are the noiseless channel outputs. N and M are, respectively, the channel and equalizer lengths. The equalizer output is denoted by $y(n) = \mathbf{w}^T(n)\mathbf{x}(n)$.

The rest of the paper is organized as follows: Section II presents the proposed criterion and related cost function; convergence properties are described in Section III; in Section IV it is shown a link between the proposed criterion and the CM one, such as Sato e Godard ones; Section V presents simulation results with the performance evaluation of the proposed criterion and, finally, in Section VI the conclusions are stated.

II. PDF Estimation-Based Blind Criterion

Let $\mathbf{w}_{\text{ideal}}$ an ideal linear equalizer, the output of the equalizer can be written as

$$y(n) = \mathbf{w}_{\text{ideal}}^T \mathbf{x}(n), \tag{1}$$

where

$$\mathbf{x}(n) = \mathbf{H}\mathbf{a}(n) + \mathbf{v}(n) \tag{2}$$

and $\mathbf{H}$ is the *convolution matrix* of the channel of dimension $M \times (N + M - 1)$ [7].

Then, using Equation (2) in (1), it is possible to write:

$$\begin{aligned} y(n) &= \left(\mathbf{H}\mathbf{a}(n) + \mathbf{v}(n)\right)^T \mathbf{w}_{\text{ideal}} \\ &= \mathbf{a}^T(n)\mathbf{H}^T\mathbf{w}_{\text{ideal}} + \mathbf{v}^T(n)\mathbf{w}_{\text{ideal}} \\ &= \mathbf{a}^T(n) \underbrace{\mathbf{H}^T\mathbf{w}_{\text{ideal}}}_{\mathbf{g}_{\text{ideal}}} + \mathbf{v}^T(n)\mathbf{w}_{\text{ideal}} \\ &= \mathbf{a}^T(n)\mathbf{g}_{\text{ideal}} + \vartheta(n) \\ &= a(n-\delta) + \vartheta(n), \end{aligned} \tag{3}$$

where $\mathbf{g}_{\text{ideal}}$ is the ideal system response, δ is a delay and $\vartheta(n)$ is a random variable (r.v.) assumed Gaussian[1] [7].

[1] This assumption is the same in Bussgang algorithms.

Equation (3) states that the pdf of the signal on the output of the equalizer is a mixture of equiprobable Gaussians (since the transmitted symbols are i.i.d.) given by:

$$p_{Y,\text{ideal}}(y) = \frac{1}{\sqrt{2\pi\sigma_\vartheta^2}} \cdot \frac{1}{S} \cdot \sum_{i=1}^{S} \exp\left[-\frac{|y(n)-a_i|^2}{2\sigma_\vartheta^2}\right], \quad (4)$$

where the a_i are the possible values of $a(n-\delta)$ that are also symbols of the transmitted alphabet $\mathcal{A}$.

Since the pdf of the equalized signal is known, it is desired to construct a criterion that forces the adaptive filter to produce signals with the same (or almost) pdf than the ideal one. This leads to the well known measure of similarities between strictly positive functions (as the pdfs), the *Kullback-Leibler Divergence* (KLD) [6].

In order to use the KLD it is constructed a parametric model, which is function of the filter parameters, to provide pdf estimation [7], a natural choice is the same model of mixture of Gaussians like that one in Equation (4) where $y(n) = \mathbf{w}^T(n)\mathbf{x}(n)$, then

$$\Phi(y, \sigma_r^2) = \underbrace{\frac{1}{\sqrt{2\pi\sigma_r^2}}}_{A} \cdot \sum_{i=1}^{S} \exp\left(-\frac{|y(n)-a_i|^2}{2\sigma_r^2}\right), \quad (5)$$

is the chosen parametric model where σ_r^2 is the variance of each Gaussian in the model.

In pattern classification field these kind of parametric functions, which are used to measure similarities with other functions, are called *target functions* [7].

Then, applying KLD to compare Equations (4) and (5) it has:

$$\begin{aligned} D_{p(y)||\Phi(y,\sigma_r^2)} &= \int_{-\infty}^{\infty} p(y) \cdot \ln\left(\frac{p(y)}{\Phi(y,\sigma_r^2)}\right) dy \\ &= \int_{-\infty}^{\infty} p(y) \cdot \ln(p(y))\, dy - \int_{-\infty}^{\infty} p(y) \cdot \ln\left(\Phi(y,\sigma_r^2)\right) dy, \end{aligned} \quad (6)$$

where $p(y) = p_{Y,\text{ideal}}(y)$.

Minimize Equation (6) is equivalent to minimize only the $\Phi\left(y, \sigma_r^2\right)$ dependent term, it means:

$$\begin{aligned} J_{\text{FP}}(\mathbf{w}) &= -E\left\{\ln\left[\Phi\left(y,\sigma_r^2\right)\right]\right\} \\ &= -E\left\{\ln\left[A \cdot \sum_{i=1}^{S} \exp\left(-\frac{|y(n)-a_i|^2}{2\sigma_r^2}\right)\right]\right\}. \end{aligned} \quad (7)$$

The ***Fitting pdf*** **(FP)** criterion corresponds to minimize $J_{\text{FP}}(\mathbf{w})$ (Equation (7)). Furthermore, it is known that minimize Equation (7) corresponds to find the entropy of y if $\Phi\left(y,\sigma_r^2\right) = p_{Y,\text{ideal}}(y)$ [8, p. 59].

An stochastic version for filter adaptation is given by:

$$\begin{aligned} \nabla J_{\text{FP}}(\mathbf{w}(n)) &= \frac{\sum_{i=1}^{S} \exp\left(-\frac{|y(n)-a_i|^2}{2\sigma_r^2}\right)(y(n)-a_i)}{\sigma_r^2 \cdot \sum_{i=1}^{S} \exp\left(-\frac{|y(n)-a_i|^2}{2\sigma_r^2}\right)} \mathbf{x}^* \\ \mathbf{w}(n+1) &= \mathbf{w}(n) - \mu_w \nabla J_{\text{FP}}(\mathbf{w}). \end{aligned} \quad (8)$$

The adaptive algorithm which uses the proposed criterion will be called ***Fitting pdf Algorithm*** **(FPA)**. Equation (8) shows an important property of the algorithm, it takes in account the phase of the transmitted symbols.

The computational complexity of this algorithm is proportional to the computation of S exponentials which are required by Equation (8). Thus, its complexity is a little higher than other LMS-like algorithms.

Other important point is, although the ideal equalizer is known to have infinity length, the use of the FP criterion does not requires a long filter to compensate the channel effect. It has been observed, through simulations, the length of the equalizer for this criterion has the same order of other blind criteria.

• *The parameter* σ_r^2*:*

As shown in the previous section, the parametric model used to update the filter coefficients is also σ_r^2 dependent. This parameter plays an important role once it is the variance of each Gaussian in the parametric model.

Moreover, σ_r^2 is also important for convergence rate because it modifies the effective step size, it means, $\mu_{\text{eff}} = \frac{\mu_w}{\sigma_r^2}$. In classification field this parameter is similar to the *temperature* one in annealing processes [7].

A numerical problem that arises with the use of the FPA is the nonconvergence for very small values of σ_r^2. This is due to the Gaussians being very sharp and much more difficult to fit the data on them. This model have also been observed in [9] where the ideal pdf of the received signal is assumed to be a mixture of impulses and later a Gaussian mixture model is considered in order to make the assumption more realistic and feasible.

Finding the optimum value for σ_r^2 parameter is still under investigation.

III. Convergence Properties

The behavior study of the FP cost function is done briefly. It is used a BPSK (Binary Phase Shift Keying) random sequence transmitted through an AR (Autoregressive) and MA (Moving Average) channels with one pole and one zero, respectively, over the real axis which have the following transfer function:

$$H_{\text{AR}}(z) = \frac{1}{1+\alpha z^{-1}}, \quad (9)$$

and

$$H_{\text{MA}}(z) = 1 + \alpha z^{-1}. \quad (10)$$

For $\alpha = 0.6$, $\mu_{\text{eff}} = 10^{-2}$, $\sigma_r^2 = 0.1$ and SNR $= 10$dB, Fig. 1 shows the cost function and the corresponding contour plot and trajectories of 20 FPA simulations (with different initializations) for the AR channel. It was considered an equalizer with two coefficients.

With the same parameters, Fig. 2 shows the cost function and the corresponding contour plot and trajectories of 20 FPA simulations (with different initializations) for the MA channel. Again, a two coefficients equalizer was considered.

One can easily note that the FP cost function presents global and local minima and convergence behavior is highly initialization dependent as in the *constant modulus* (CM) criterion. This characteristic shows (as Figs. 1 and 2 illustrate) the phenomenon

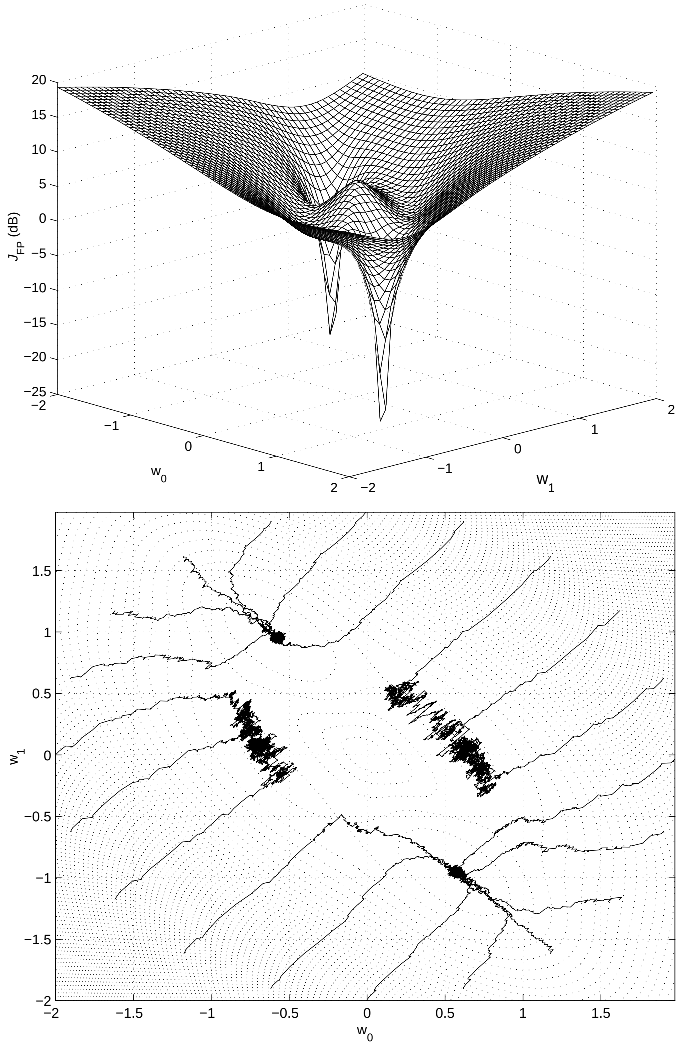

Fig. 1. (Up) FP cost function for AR channel. (Down) Corresponding contour plot and trajectories of 20 simulations ($\mu_{\text{eff}} = 10^{-2},\ \sigma_r^2 = 0.1$).

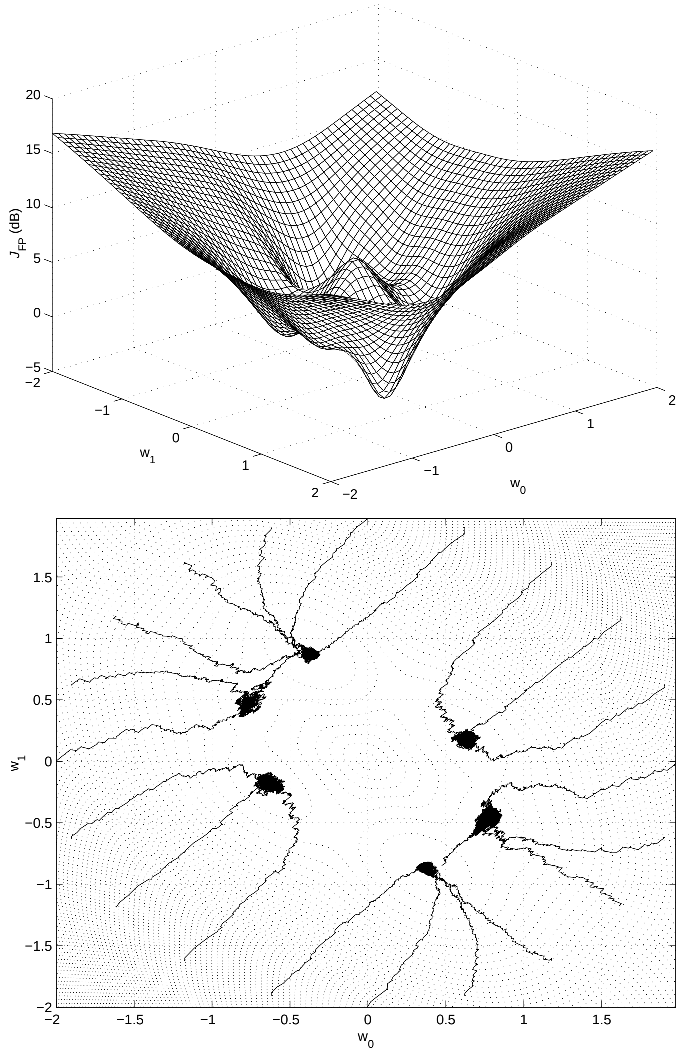

Fig. 2. (Up) FP cost function for MA channel. (Down) Corresponding contour plot and trajectories of 20 simulations ($\mu_{\text{eff}} = 10^{-2},\ \sigma_r^2 = 0.1$).

of ill-convergence also present in the FP criterion. It worths to mention that the cost function for the CM criterion is pretty similar to the FP one. Nevertheless the presence of local minima in the cost function, it has been observed that the global minima are pretty close to those ones from the Wiener solution [10].

Next section is devoted to formalize mathematically the local convergence property of the criterion observed in the simulations.

• *Convergence in FP Criterion:*

Inspired by the local convergence analysis presented in [11], which uses a mixed adaptation strategy switching from the CMA to a criterion based on a mixture of Gaussians, in this section is presented a brief analysis that shows the local convergence property observed by computational simulations.

As previously stated (Section II), the equalizer inputs are given by $\mathbf{x} = \mathbf{H}\mathbf{a}(n) + \mathbf{v}(n)$. The aim is to obtain a filter as close as possible of an optimum one represented by $\mathbf{w}_{\text{opt}}$ which is considered to provide $\mathbf{w}_{\text{opt}}^T\mathbf{H}\mathbf{a}(n) = a(n-\delta)$ where $a(n-\delta) = a_i$ and $1 \leq i \leq S$. It is also considered some perturbation in the optimum filter estimation, it means $\mathbf{w} = \mathbf{w}_{\text{opt}} + \Delta\mathbf{w}$, in order to analyze its influence on the convergence when the solution is near to a region of minimum (global or local).

In the cost function (7) it is retained only the term corresponding to $a_k = a(n-\delta)$ which the equalizer output is closer and the other components are neglected. Thus,

$$\begin{aligned}
&J_{\text{FP}}(\mathbf{w}) = \\
&= -E\left\{\ln\left[A \cdot \sum_{i=1}^{S} \exp\left(-\frac{\left|\mathbf{w}^T\mathbf{H}^T\mathbf{a}(n) + \mathbf{w}^T\mathbf{v}(n) - a_i\right|^2}{2\sigma_r^2}\right)\right]\right\} \\
&\approx -E\left\{\ln\left[A \cdot \exp\left(-\frac{\left|\Delta\mathbf{w}^T\mathbf{H}^T\mathbf{a}(n) + \mathbf{w}^T\mathbf{v}(n)\right|^2}{2\sigma_r^2}\right)\right]\right\} \\
&\approx -E\left\{-\frac{\left|\Delta\mathbf{w}^T\mathbf{H}^T\mathbf{a}(n) + \mathbf{w}^T\mathbf{v}(n)\right|^2}{2\sigma_r^2} + \ln(A)\right\} \\
&\approx \frac{\sigma_a^2\Delta\mathbf{w}^H\mathbf{H}\mathbf{H}^H\Delta\mathbf{w} + \sigma_n^2\mathbf{w}^H\mathbf{w}}{2\sigma_r^2} + \ln(A)
\end{aligned} \tag{11}$$

where σ_a^2 indicates the average power of the transmitted sequence and σ_n^2 is the input noise variance. Hence, the gradient can also be approximate by:

$$\nabla J_{\text{FP}}(\mathbf{w}) \approx \frac{\sigma_a^2}{2\sigma_r^2}\left(\mathbf{H}\mathbf{H}^H \Delta\mathbf{w} + \frac{\sigma_n^2}{\sigma_a^2}\mathbf{w}\right) \tag{12}$$

where $\frac{\sigma_n^2}{\sigma_a^2}$ is the inverse of the signal-to-noise ratio (SNR).

In the absence of noise (SNR $\rightarrow \infty$) the following LMS adaptation rule is obtained:

$$\begin{aligned} \mathbf{w}(n+1) &= \mathbf{w}(n) - \mu\frac{\sigma_a^2}{\sigma_r^2}\mathbf{H}\mathbf{H}^H \Delta\mathbf{w}(n) \\ \mathbf{w}_{\text{opt}}(n+1) &+ \Delta\mathbf{w}(n+1) = \mathbf{w}_{\text{opt}}(n) + \Delta\mathbf{w}(n) \\ &- \mu\frac{\sigma_a^2}{\sigma_r^2}\mathbf{H}\mathbf{H}^H \Delta\mathbf{w}(n) \end{aligned} \tag{13}$$

Since the optimum weight vector is not time instant-dependent it follows:

$$\Delta\mathbf{w}(n+1) = \Delta\mathbf{w}(n)\left(\mathbf{I} - \mu\frac{\sigma_a^2}{\sigma_r^2}\mathbf{H}\mathbf{H}^H\right), \tag{14}$$

where $\mathbf{I}$ is the identity matrix. This recursive rule converges if:

$$0 < \mu < \frac{4\sigma_r^2}{\sigma_a^2 \lambda_{\max}} \tag{15}$$

where $\lambda_{\max}$ is the greatest eigenvalue of $\mathbf{H}\mathbf{H}^H$.

The result in (15) confirms our intuition, stated in Section II, that the σ_r^2 strongly influences the convergence speed due to its "control" on the convergence factor.

Unfortunately, it cannot be assured the global convergence independently of the initialization as in the CM criterion. This issue is still under study for a formal mathematical formulation.

The results presented in this section could indicate that CM and FP criteria have a link and this issue is addressed in next section.

IV. A Link to the CM Criterion

In the previous section it has been observed that the FP criterion has the same convergence characteristics than the CM one, then it seems reasonable to compare the FP criterion with it and find some link or equivalence.

Thinking about measure of similarities between functions, other criteria have also a parametric function (target function, $f_{\text{CM}}(\cdot)$) used to find the equalizer outputs. In that case, the CM criterion is written as a measure of similarities between functions to find some link between both criteria, the CM and FP ones.

Using the same approach to find the cost function, for the CM criterion the following equation is given:

$$J_{\text{CM}}(\mathbf{w}) = -E\left\{\ln\left(f_{\text{CM}}\left(y, \mathbf{w}, p\right)\right)\right\}, \tag{16}$$

where p chooses the criterion to Sato criterion ($p = 1$) and Godard one ($p = 2$). But it is known that

$$J_{\text{CM}}(\mathbf{w}) = E\left\{\left(|y|^p - R_p\right)^2\right\}, \tag{17}$$

then to find the equivalent target function, Equations (16) and (17) have to be equal, resulting

$$f_{\text{CM}}(y, \mathbf{w}, p) = \frac{\exp\left[\left(|y|^p - R_p\right)^2\right]}{\int_{-\infty}^{\infty} \exp\left(|\xi|^p - 1\right) d\xi}, \tag{18}$$

where the term in the denominator is used to guarantee that $f_{\text{CM}}(y, \mathbf{w}, p)$ has unity module. It can be then seen that the CM criterion has the same structure for the parametric function that the FP one.

It has also been simulated the target functions for the FPC with different values of σ_r^2 and compared them with those ones from the CM criterion (Sato and Godard), Fig. 3 shows the target functions for BPSK modulation.

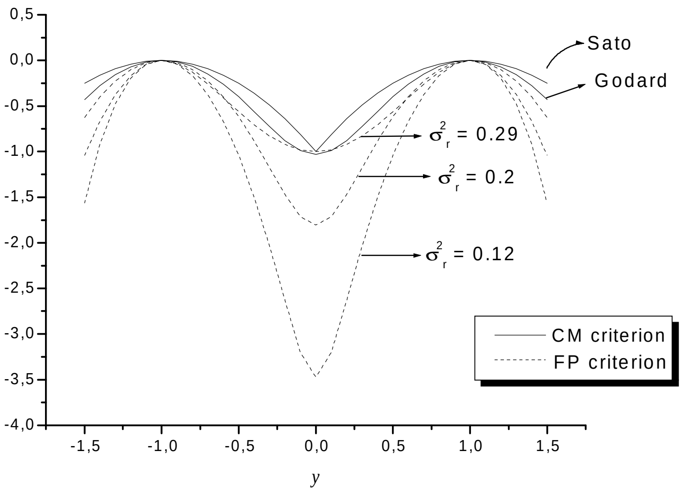

Fig. 3. Target functions for FP and CM criteria.

It is observed that for some choice of σ_r^2 the CM and FP target functions are equivalent. It wonders that this value is modulation type-dependent and, as mentioned in Section II, an strategy to find the optimum value is still under investigation.

Moreover, for more complex modulation schemes, the FP criterion target function can better fits the idealized equalizer output pdf with S Gaussian kernels. Thus, since Equation (8) can take into account even complex symbol parts and multilevel alphabets, it can be predicted an improved performance w.r.t. the CM criterion.

V. Simulation Results

For performance evaluation, the FPA has been compared with the CMA in the equalization context. A QPSK modulation transmitted over a channel with the following impulse response is used:

$$\mathbf{h} = [\, 2 - \text{j}0.4 \quad 1.5 + \text{j}1.8 \quad 1 \quad 1.2 - \text{j}1.3 \quad 0.8 + \text{j}1.6 \,]^T. \tag{19}$$

This channel has been used in some recent works to illustrate the robustness of algorithms facing a very distorsive channel [12].

As merit figure, it has been used the measure of residual interference (RI), defined as:

$$\mathrm{RI}(n) = \frac{\left| \sum_k |g_k(n)| - \max_k |g_k(n)| \right|}{\max_k |g_k(n)|}, \qquad (20)$$

where $g_k(n)$ is the k-th element of the vector $\mathbf{g}(n) = \mathbf{H}^T\mathbf{w}(n)$.

The simulation parameters were: filters with 30 taps for both algorithms with center-spiked initialization [6] and $\mu_w = 10^{-3}$ for both algorithms and SNR = 30 dB. FPA has also used $\sigma_r^2 = 0.29$. Fig. 4 shows the RI evaluation for both algorithms.

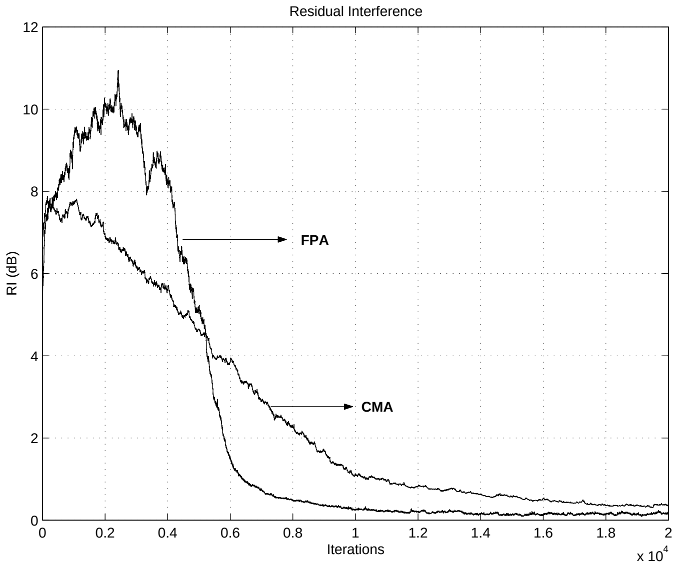

Fig. 4. Residual interference evolution for FPA and CMA.

One can easily note that FPA has a faster convergence than the CMA, it also worths to mention is that final performance is practically the same for both of them with a little gain from the FPA. The faster convergence is due to the fact FPA uses more suitable target function that better fits the transmitted signals characteristics than the CMA that considers the modulus only. However, it should be strongly emphasize that CMA requires a phase recovering device for correctly match the constellation with the transmitted one whereas the FPA could done it by itself, as shown in Figures 5 and 6.

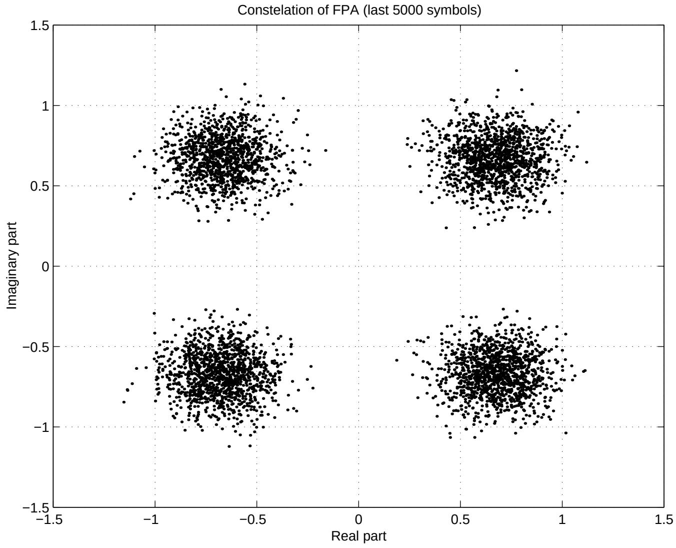

Fig. 5. Equalizer output signal constellation for the FPA.

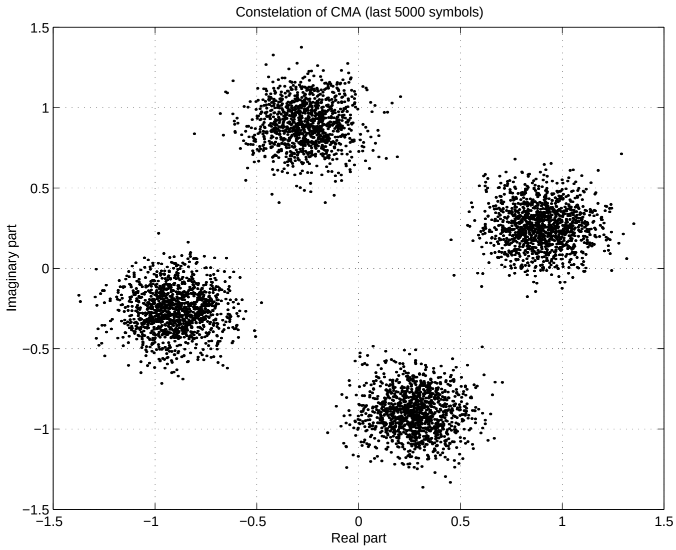

Fig. 6. Equalizer output signal constellation for the CMA.

VI. Conclusions and Perspectives

In this paper it was proposed a new blind information-theoretic based criterion for adaptive equalization. The derivation of the cost function is done by means of the estimation of the ideally equalized signal through a parametric model resulting in a sort of entropy minimization of the equalizer output signal. Convergence properties are briefly studied in order to clarify some important points.

It has also been shown that the proposition has a link with the constant modulus criterion when that criterion is studied as a measure of similarities between two functions.

Through simulations it has been observed that, for some cases, the FPA presents faster convergence than the CMA, this property and the interesting issue of the phase recover ability of the FPA may indicate a plausible alternative for blind equalization.

As perspectives to this work, we note that the parameter σ_r^2 plays an important role on the FP cost function and deserves much more investigation on it. As well, the consideration of more complex modulations (nonconstant modulus) may provide some gain to the FPA once it considers a target function with the number of Gaussian kernels equals the number of symbols from the modulation alphabet. As another front, a multiuser version of the algorithm is under investigation for space-time signal detection.

Acknowledgements

The authors would like to thank Romis R. F. Attux by his useful comments on the convergence analysis issue.

REFERENCES

[1] Y. Sato, "A Method of Self-recovering Equalization for Multi-level Amplitude Modulation," *IEEE Trans. on Communications*, vol. 23, pp. 679–682, June 1975.

[2] D. N. Godard, "Self-Recovering Equalization and Carrier Tracking in Two-Dimensional Data Communication Systems," *IEEE Trans. on Communications*, vol. COM-28, pp. 1867–1875, November 1980.

[3] S. Haykin, ed., *Blind Deconvolution*. (Information and Systems Sciences Series), NJ, USA: Prentice-Hall, Englewood Cliffs, 1994.

[4] A. J. Bell and T. Sejnowski, "An Information-Maximisation Approach to Blind Separation and Blind Deconvolution," *Neural Computation*, vol. 7, pp. 1129–1159, November 1995.

[5] J. Sala-Alvarez and G. Vázquez-Grau, "Statistical Reference Criteria for Adaptive Signal Processing in Digital Communications," *IEEE Trans. on Signal Processing*, vol. 45, pp. 14–31, January 1997.

[6] S. Haykin, ed., *Unsupervised Adaptive Filtering*, vol. II of *Series on Adaptive and Learning Systems for Signal Processing, Communications and Control*. John Wiley & Sons, blind deconvolution ed., 2000.

[7] C. C. Cavalcante, "Neural prediction and probability density function estimation applied to blind equalization," Master's thesis, Federal University of Ceará, February 2001.

[8] C. M. Bishop, *Neural Networks for Pattern Recognition*. UK: Oxford University Press, 1995.

[9] R. Amara, *Égalisation de Canaux Linéaires et Non Linéaires, Approche Bayésienne*. Ph.D. thesis (in French), Université Paris XI Orsay, France, November 2001.

[10] S. Haykin, *Adaptive Filter Theory*. (Prentice-Hall Information and System Sciences Series), Prentice-Hall, 3^{rd} ed., 1996.

[11] S. Barbarossa and A. Scaglione, "Blind Equalization Using Cost Function Matched to the Signal Constelation," in *Proc. 31^{st} Asilomar Conference on Signals, Systems and Computer*, (Pacific Grove (CA) - USA), pp. 550–554, November 1997.

[12] J. Labat, O. Macchi, and C. Laot, "Adaptive Decision Feedback Equalization: Can You Skip the Training Period?," *IEEE Trans. on Communications*, vol. 46, pp. 921–930, July 1998.