A LINEARLY CONSTRAINED BLIND EQUALIZATION SCHEME BASED ON BUSSGANG TYPE ALGORITHMS

Santiago Zazo; J.M. Páez-Borrallo; Iván A. Pérez Alvarez**
*Universidad Alfonso X El Sabio. Madrid; * ETS Ingenieros de Telecomunicación - Universidad Politécnica de Madrid*
Phone: 341-3367280; Fax: 341- 3367350; e-mail: santiago@gaps.ssr.upm.es

ABSTRACT

Existing blind adaptive equalizers that use nonconvex cost functions (as Bussgang type algorithms) and stochastic gradient descent suffer from lack of global convergence to an equalizer tap set that removes sufficient ISI when an FIR equalizer is used. In this paper we propose a new algorithm including a tap anchoring and gain recovering into the classical schemes. The combined effect of these strategies is to establish the preservation of the transmitted symbol preventing for ill convergence, and therefore providing the ability of implementation the inverse filter regardless of the initial ISI. Under certain hypotheses, we suggest that a globally convex scheme can be proposed overcoming the existing structures. Several computer simulations support our theoretical results.

1. INTRODUCTION

Blind adaptive channel equalizers are important devices in high data rate, bandlimited digital communication systems in which the transmission of a training sequence is impractical or very costly. In this way, the receiver can begin its self adaptation without the assistance of the transmitter. The ability of blind start-up also enables a blind equalizer to self recover from system breakdowns, during which the equalizer may have lost track of the desired parameter settings. Many techniques for blind equalization have been proposed in the recent literature, which the most important are the Bussgang type algorithms and the higher order statistics (HOS) based algorithms [1].

Bussgang type algorithms intend the estimation of the transmitted symbol by means of a zero memory non linearity, as an internally desired response generator. However, the main drawback of these schemes is the lack of convexity, where any stochastic gradient algorithm can be trapped in a local minimum, sometimes far away of the desired response [1].

On the other hand, methods exist in which the channel coefficients are obtained by solving systems of nonlinear equations that involve the higher order cumulants of the channel output [1,2]. Some recursive algorithms based on cumulants have also been reported but they are computationally expensive. It must be remarked that the main drawback to-date of these methods is that very long data lengths are needed in order to reduce the variance associated with estimating the higher order statistics from real data using sample-averaging techniques. Also, these methods trying to estimate the inverse filter neglecting the noise effect, fall down when the channel has spectral nulls because the noise enhancement may degrade critically the desired behavior.

Therefore, Bussgang type blind equalizers should be very attractive, if we were able to avoid the presence of local minima: in this line, techniques based on sopt&go schemes have been proposed providing effective blind convergence in the MSE sense [3]: a simple flag tells the equalizer whether the current output error is sufficientlly reliable to be used; in this way, it works retaining the advantages of simplicity of the Decision Directed Algorithm (DDA) while attempting to substantially improve its blind convergence capabilities. However, it must be pointed out that the cost function remains multimodal.

Another research line has been developed in the recent years in order to find new cost function achieving unimodality by introducing a linearly constrained equalizer with a single constrained tap [4, 5]. The methods proposed are based on the original work by Lucky [6] who reported the strategy of tap anchoring and also analyzed its effect on the global convexity of the ISI criterion function he considered. However, their main drawback is the slow convergence of the gradient descent approaches.

Also, tap anchoring has been the starting point of our research [7, 8]: our proposal consist in a linearly constrained equalizer also with an adaptive zero memory nonlinearity for gain /phase tracking. In this work we present the scheme supported by an important theoretical background showing its globally convex characteristic for non minimum phase channels, under certain hypotheses about the ISI distribution. In particular we have developed an algorithm labeled as Modified Decision Directed Algorithm (MDDA), which is basically a modification to the standard DDA, including a fixed tap equalizer and an adjustable decision device. Additionally, this scheme can be easily extended to other Bussgang type algorithms as the Constant Modulus Algorithm (CMA). The main idea in our opinion about the tap anchoring, is the fact that this strategy preserves in some way the transmitted symbol and therefore prevents for ill convergence.

2. THE MDDA

Let us start showing the block diagram of the MDDA in Fig.1: an unknown PAM sequence is transmitted (the extension to QAM is straightforward) through a generic non minimum phase channel delaying Δ samples and also introducing a noisy effect. In the receiver, a two sided equalizer ($2N$+1 taps) intends the ISI elimination. Also, it introduces a blind automatic gain control (AGC) to estimate and compensate the channel amplitude. Finally, a decision device including an adjustable gain (denoted α) estimates

This work has been supported by TIC92-0329

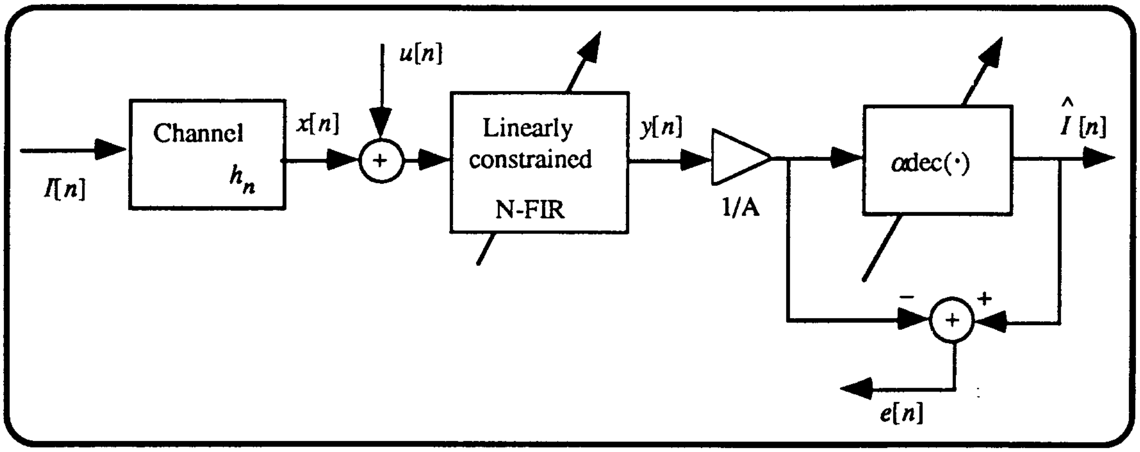

Fig.1: Block diagram of the MDDA

the transmitted symbol also driving the error measure $e[n]$.

First of all, let us express the equalizer output as:

$$y = \mathbf{c}^t \mathbf{H} \mathbf{I} \tag{1}$$

where $\mathbf{c}$ is the equalizer coefficient vector, $\mathbf{H}$ is the channel convolutional matrix and $\mathbf{I}$ is the transmission vector. Splitting the terms in (1) into the current symbol and the remainder, and also remarking the special role of the center tap we have:

$$\mathbf{c} = \begin{pmatrix} \mathbf{c}_a \\ c_N \\ \mathbf{c}_b \end{pmatrix} \to \begin{pmatrix} c_N \\ \mathbf{c}_1 \end{pmatrix} \qquad \mathbf{I} = \begin{pmatrix} \mathbf{I}_a \\ I \\ \mathbf{I}_b \end{pmatrix} \to \begin{pmatrix} I \\ \mathbf{I}_1 \end{pmatrix} \tag{2}$$

where $\mathbf{c}_1^t = \begin{pmatrix} \mathbf{c}_a^t & \mathbf{c}_b^t \end{pmatrix}$ and $\mathbf{I}_1^t = \begin{pmatrix} \mathbf{I}_a^t & \mathbf{I}_b^t \end{pmatrix}$

$$\mathbf{H} = \begin{pmatrix} \ddots & \vdots & \vdots & \vdots & \\ \cdots & h_\Delta & h_{\Delta+1} & h_{\Delta+2} & \cdots \\ \cdots & h_{\Delta-1} & h_\Delta & h_{\Delta+1} & \cdots \\ \cdots & h_{\Delta-2} & h_{\Delta-1} & h_\Delta & \cdots \\ & \vdots & \vdots & \vdots & \ddots \end{pmatrix} = \begin{pmatrix} \mathbf{H}_a & \mathbf{h}_a & \mathbf{H}_b \\ \mathbf{h}_b^t & h_\Delta & \mathbf{h}_c^t \\ \mathbf{H}_d & \mathbf{h}_d & \mathbf{H}_c \end{pmatrix} \to$$

$$\to \begin{pmatrix} h_\Delta & \mathbf{h}_1^t \\ \mathbf{h}_2 & \mathbf{H}_1 \end{pmatrix} \text{ where } \mathbf{H}_1 = \begin{pmatrix} \mathbf{H}_a & \mathbf{H}_b \\ \mathbf{H}_d & \mathbf{H}_c \end{pmatrix}, \tag{3}$$

$$\mathbf{h}_1^t = \begin{pmatrix} \mathbf{h}_b^t & \mathbf{h}_c^t \end{pmatrix}, \quad \mathbf{h}_2 = \begin{pmatrix} \mathbf{h}_a & \mathbf{h}_d \end{pmatrix}$$

After some algebra we reach the following equation:

$$y = c_N h_\Delta I + \mathbf{c}_1{}^t \mathbf{h}_2 I + c_N \mathbf{h}_1^t \mathbf{I}_1 + \mathbf{c}_1{}^t \mathbf{H}_1 \mathbf{I}_1 \tag{4}$$

As we proposed in [7] for minimum phase channels, a good criterium to design the linear constraint is to preserve always the transmitted symbol (the remainder terms are considered as ISI), i.e.

$$c_N h_\Delta + \mathbf{c}_1{}^t \mathbf{h}_2 \neq 0 \tag{5}$$

However, the choice for no minimum phase channels is not so simple as we proposed because the current symbol depends on the whole equalizer parameter set; in the minimum (maximum) phase case it can be shown that for a one side equalizer, the current symbol depends only on the first (last) tap, and so that, the obvious choice is to fix this tap leading to a linear forward (backward) prediction problem. In a generic case given by eq. (4), it could be a good idea considering a practical constraint to fix the center tap, assuming that the channel impulse response is maximum for a Δ samples delay (this intuitive choice connects with the early work by Lucky[6] and the anchoring tap in [4, 5 and references therein]). Also in this line we want to recall the work developed in [9], where a linearly constrained filter is considered as the convolution of a forward and a backward linear predictors, each one implementing the minimum and maximum phase which every nonminimum phase channel can be decomposed.

3. MDDA. ANALYSIS OF CONVEXITY.

The error surface proposed is given by:

$$J = E\left\{ \left(y - \alpha \mathrm{dec}(y) \right)^2 \right\} \tag{6}$$

Taking derivatives, the gradient vector yields:

$$\frac{\partial J}{\partial \mathbf{c}_1} = 2E\left\{ y \frac{\partial y}{\partial \mathbf{c}_1} \right\} - 2\alpha E\left\{ \mathrm{dec}(y) \frac{\partial y}{\partial \mathbf{c}_1} \right\} \tag{7a}$$

$$\frac{\partial J}{\partial \alpha} = 2\alpha E\left\{ \mathrm{dec}^2(y) \right\} - 2E\left\{ y \mathrm{dec}(y) \right\} \tag{7b}$$

Substituting (7b) into (7a), the condition of stable points yields:

$$2E\left\{ y \frac{\partial y}{\partial \mathbf{c}_1} \right\} - 2E\left\{ y \mathrm{dec}(y) \right\} E\left\{ \mathrm{dec}(y) \frac{\partial y}{\partial \mathbf{c}_1} \right\} = 0$$
$$\text{where } \frac{\partial y}{\partial \mathbf{c}_1} = \mathbf{h}_2 I + \mathbf{H}_1 \mathbf{I}_1 \tag{8}$$

After some algebra we obtain:

$$\mathbf{h}_2\left(E\{yI\} - E\{y\mathrm{dec}(y)\} E\{I\mathrm{dec}(y)\} \right) +$$
$$E\{\mathbf{H}_1 \mathbf{I}_1 y\} - E\{y\mathrm{dec}(y)\} E\{\mathbf{H}_1 \mathbf{I}_1 \mathrm{dec}(y)\} = 0 \tag{9}$$

The main problem of this analysis is the evaluation of the expected values involving the thresholding function and vector $\mathbf{H}_1\mathbf{I}_1$; in [10] is suggested a useful relation for the calculus of cross-moments between zero mean gaussian random variables involving the sign function. Following that idea, recall that assuming a conditioned model for the residual ISI, **Y** conditioned to **I**=I can be considered as a gaussian non zero mean random variable [11]. Also, notice that each component of vector $\mathbf{H}_1\mathbf{I}_1$ (in the sequel we will denote as vector **x**), is obtained as a linear combination of many i.i.d. random variables, and applying the Central Limit Theorem, the gaussian hypothesis is supported so much the longer the channel is; summarizing, we propose the following hypotheses:

$$\begin{array}{ll} \mathbf{Y}|I & N(c_N h_\Delta I + \mathbf{c}_1{}'\mathbf{h}_2 I, \sigma_\varepsilon) \\ \mathbf{x} & N(0, \sigma_\mathbf{x}) \end{array} \qquad (10)$$

where ε means ISI.

However, we can not apply directly the expression given in [10], because our random variables although gaussian are non zero mean. So that, we have derived a generalization of the expression in [10] for generic gaussian random variables.

Theorem 1: Let **X** and **Y** as gaussian random variables $N(0,\sigma_\mathbf{x})$ and $N(\eta_\mathbf{y},\sigma_\mathbf{y})$ respectively and ρ as the correlation coefficient. Also, let us denote g(·) as a generic non linear transformation; it can be shown that:

$$E\{xg(y)\} = \frac{E\{xy\}E\{(y-\eta_y)g(y)\}}{E\{(y-\eta_y)^2\}} \qquad (11)$$

Applying theorem 1, (9) can be expressed as:

$$\mathbf{h}_2\left(E\{yI\} - \frac{E\{y\mathrm{dec}(y)\}}{E\{\mathrm{dec}^2(y)\}}E\{I\mathrm{dec}(y)\}\right) + E\{\mathbf{xy}\}\left(1 - \frac{E\{y\mathrm{dec}(y)\}}{E\{\mathrm{dec}^2(y)\}}\frac{E\{(y-\eta_y)\mathrm{dec}(y)\}}{E\{(y-\eta_y)^2\}}\right) = 0 \qquad (12)$$

Note that (12) is a vectorial equation where each component of both terms are multiplied by two scalar functions depending on the ISI power and a gain factor (recall (4) where the equalizer output can be considered as a composition of the scaled current symbol and a gaussian ISI: Y=aI+ε)

$$\mathbf{h}_2 g_1(y) + E\{\mathbf{xy}\}g_2(y) = 0 \qquad (13)$$

In [8] we have reached a very similar conclusion but for binary transmission (or 4QAM). We concluded there that if we were able to preserve the transmitted symbol (anchoring the center tap yields as sufficient condition for several channels we have tried), we assured that there is only one stable point. However, for higher order constellations there is a critical point in the gain recovery; we have observed that even for minimum phase channels, if the gain is not recovered we can not assure the right convergence. Therefore, assuming that the AGC works fairly well, and so on the coefficient of the current symbol is approximately 1, functions g_1 and g_2 have been depicted in Fig.2. Here, we want to point out that if this hypothesis is not assumed, we can not guarantee the global characteristic of the cost function proposed. However, simulations support our analysis.

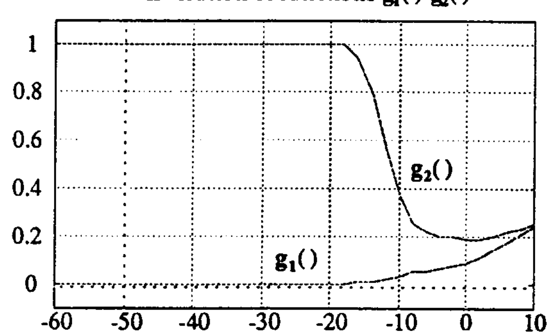

Fig.2 Evolution of functions g_1 and g_2

Solving equation (13) to find the stability points, let us consider the following cases:

a) For advanced convergence levels, first factor vanishes. Since $g_2(y)$ never vanishes, the stable point arises from the correlation factor, i.e.:

$$E\{\mathbf{xy}\} = 0 \qquad (14)$$

b) For not advanced convergence levels, neither $g_1(y)$ or $g_2(y)$ vanish. So that, the stable points must keep the following condition for each component:

$$E\{\mathbf{xy}\} \propto \mathbf{h}_2 \qquad (15)$$

Recall that (15) never can happen because each component of the cross-correlation depends on the whole channel impulse response, meanwhile, each component of the other term is just one sample of the channel impulse response. So that (6) has only one stable point, solving equation (14), that is:

$$\mathbf{c}_{1_{opt}} = -c_N(\mathbf{HH}')^{-1}\mathbf{H}_1\mathbf{h}_2 \qquad \alpha_{opt} = \left.\frac{E\{y\mathrm{dec}(y)\}}{E\{\mathrm{dec}^2(y)\}}\right|_{\mathbf{c}_1 = \mathbf{c}_{1_{opt}}} \qquad (16)$$

4.LCCMA WITH ADAPTIVE PARAMETER. ANALYSIS OF CONVEXITY.

Once we have developed the analysis of convexity for the MDDA, the results are easily extended to other schemes as the Linear Constrained Constant Modulus Algorithm (LCCMA) with variable parameter. This method was presented in beamforming applications [12 and references therein] as a combination of two existing untrained adaptive techniques: linearly constrained minimization of output power and unconstrained minimization of complex envelope or modulus variations. Additionally, it is pointed out the possibility of making the CMA parameter also adaptive. In this case, the cost error function is:

$$J = E\left\{\left[|y|^2 - \gamma\right]^2\right\} \qquad (17)$$

Taking derivatives in (17) we find the gradient:

$$\frac{\partial J}{\partial c_1} = 2E\left\{ yy^* \frac{\partial y}{\partial c_1} y^* \right\} - 2\gamma E\left\{ \frac{\partial y}{\partial c_1} y^* \right\} \quad (18a)$$

$$\frac{\partial J}{\partial \gamma} = 2\gamma - 2E\left\{ yy^* \right\} \quad (18b)$$

and substituting (18b) into (18a) and after some algebra, we obtain the same stability condition given by (13); Additionally, it is much more simpler because in this case, functions g_1 and g_2 are given by:

$$\begin{cases} g_1(y) = a\sigma_\varepsilon^2 \\ g_2(y) = a^2 \end{cases} \quad (19)$$

At this point, we follow the same argument to show that this scheme is globally convex also with only one minimum whose position is given by (16).

5. COMPUTER SIMULATIONS

Finally, some computer simulations support our proposal. We have chosen a typical telephone channel impulse response [13]:

$$H(z) = 0.04 - 0.05z^{-1} + 0.07z^{-2} - 0.21z^{-3} - 0.5z^{-4}$$
$$+0.72z^{-5} + 0.36z^{-6} + 0.21z^{-8} + 0.03z^{-9} + 0.07z^{-10} \quad (20)$$

First at all, we will like to show that (16) is in fact the optimum equalizer. We have implemented equation (16) for the channel given by (20) for a 31 taps equalizer; the impulse response is given in Fig.3 (the convolution of the channel and equation (16) is the Kronecker delta).

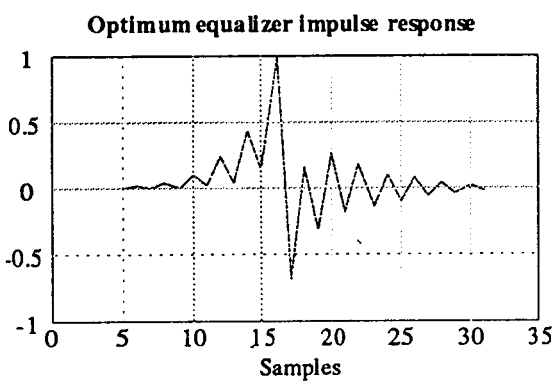

Fig.3. Optimum equalizer impulse response

To show the ability of the methods proposed to recover the transmitted symbols, we have implemented both instantaneous stochastic gradient descent algorithms. In order to prove its global convergent characteristic, we have chosen the initialization point randomly observing that always is reached the right solution. We have considered a 16QAM transmission for the MDDA, and a 8PSK for the MCMA with a 31 taps equalizer. See Fig.4, 5.

5. REFERENCES

[1] Haykin, S.*Blind deconvolution..* Englewood Cliffs, Prentice Hall, 1994.

[2]Nikias,C.,J.Mendel. Signal processing with Higher Order Spectra. *IEEE Signal Processing Magazine*, pp.10-37,July 1993.

[3] Picchi, Giorgio,Giancarlo Pratti. Blind equalization and carrier recovery using a "stop-and-go" decision directed algorithm. *IEEE Transactions on Communications* vol. COM-35, No9, pp.877-887. September 1987

[4] Yamazaki, K., R.A. Kennedy. Reformulation of linearly constrained adaptation and its application to blind equalization. *IEEE Transactions on Signal Processing*, Vol.42, No7, pp. 1837-1841, July 1994.

[5] Vembu, S., S. Verdú, R.A. Kennedy, W. Sethares. Convex cost functions in blind equalization. *IEEE Transactions on Signal Processing,*. Vol.42, No8, pp. 1952-1959, August 1994.

[6] Lucky, R.W. Automatic equalization for digital communication. The Bell System T echnical Journal, April 1965.

[7] Zazo,S., J.M. Páez Borrallo, I.A. Pérez Alvarez. A globally convex error surface in a modified decision directed algorithm for blind equalization. COST #229. pp. 122-128, Bayona (Vigo) Spain, June, 1993.

[8] Zazo,S., J.M. Páez Borrallo, I.A. Pérez Alvarez. Analysis of a Globally Convex Algorithm for Blind Equalization of Non Minimum Phase Channels. *EUSIPCO 94*, Edinburg, U.K. 1994.

[9] Rocha, Faria da, Odile Macchi. A novel self learning adaptive recursive equalizer with unique optimum for QAM. *ICASSP* 94 pp. III.481-III.484, Adelaida 1994.

[10] V. J. Mathews, S. H. Cho: Improved Convergence Analysis of Sthocastic Gradient Adaptive Filters using the Sign Algorithm.*IEEE Transactions on Acoustics, Speech and Signal Processing,* vol.ASSP-35, No.4,pp. 450-454, April 1987.

[11] Shynk, J.J., C.K. Chan: Performance surfaces of the constant modulus algorithm based on a conditional Gaussian Model. *IEEE Transactionson Signal Processing,*. Vol.41, No5, pp. 1965-1969, May 1993.

[12] Rude, M.J. A linearly constrained adaptive algorithm for constant modulus signal processing. Ph. D. Thesis, University of Southern California, August 1990.

[13] Proakis, J.G. Digital Communications, Singapore, Mac Graw Hill, 1989.

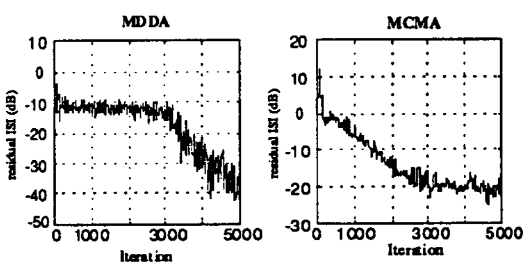

Fig.4 Learning curves for the MDDA and MCMA.

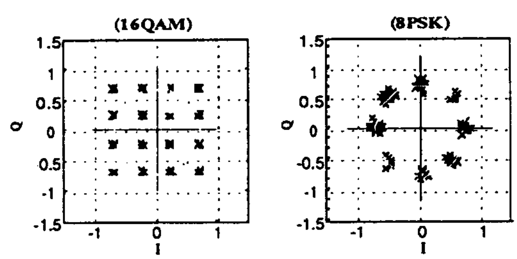

Fig.5 Eye pattern at 5000