CONVEXITY IN SISO BLIND EQUALIZATION

Mónica Corlay [*], *Maurice Charbit*

ENST/TSI
46 rue Barrault
75634 Paris Cedex, France.
(corlay,charbit)@tsi.enst.fr

Pierre Duhamel

CNRS/LSS
Plateau du Moulon
F-91192 Gif sur Yvette Cedex, France.
pierre.duhamel@lss.supelec.fr

ABSTRACT

In [1], a new class of blind equalization cost functions was proposed. They have the particularity of being unimodal, and some of them, sectionally convex in the Combined System Domain (CDS), both properties seeming very attractive for a cost function, though no proof (or specific definition) was given of the sectional convexity. In this paper we show, using a particular case of this class, that sectional convexity not only implies to fix a delay, but that it also requires to fix the value of the coefficient associated to such delay. We show that the CMA criterion shares the same property, and the difficulty of maintaining this property in actual algorithms.

1. INTRODUCTION

The problem of blind equalization (in the noiseless case) is illustrated in figure (1). The idea is to retrieve the input sequence $\{a_n\}$ sent through an unknown channel (with impulse response coefficients $\{h_i\}$), with the aid of the channel output signal $\{y_n\}$, a linear equalizer (with impulse response coefficients $\{w_i\}$) and some knowledge of the statistics of the original sequence. The criterion usually involves higher order statistics (HOS) of the emitted sequence, but not the sequence itself. The use of HOS is justified because it is well known that second order statistics are insufficient to equalize non minimum phase channels. Since [2] many

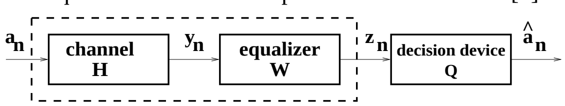


Fig. 1. Blind Equalization Model

adaptive blind equalization algorithms have been proposed to estimate $\{w_i\}$ such as [3] (or CMA [4]) among others. They are computationally simple and based on the minimization of a cost function by a stochastic gradient algorithm. However, for most cost functions, unimodality has not been proven in the situations met in practice (finite length equalizers). This can lead to possible misconvergence and insufficient removal of the Intersymbol Interference (ISI) if the algorithm converges to a stable local minimum. The problem of the existence of local minima (more specifically, those of the CMA) has been the subject of many papers such as [5] [6] that study the problem in the noiseless case. In this paper we consider only the single channel in the noiseless case, since the single input multiple output case is very different in nature.

[*] Author supported by CONACYT, México.

2. UNIMODALITY VS. CONVEXITY

The properties of classical cost functions are easily obtained in terms of the input signal $\{a_n\}$ of the total system, since they are usually based on the property that the emitted sequence is i.i.d. Hence the results involve the global transfer function (i.e. the combined system). We want to know how the properties that are obtained in terms of the CSD extend to the equalizer.

The impulse response of the combined channel-equalizer system shown on figure(1), denoted as $c_i = \sum_j h_{i-j} w_j$, defines a linear transformation from the equalizer domain (ED) to the CSD. We note this transformation $C = \mathcal{H}^T W$, where $\mathcal{H}$ is a linear operator. We have perfect equalization when $ISI = 0$, i.e. when there exists an arbitrary delay d such that, in the CSD:

$$c_d \neq 0, \quad \text{and} \quad \forall i \neq d, \, c_i = 0 \tag{1}$$

Note that the condition is defined up to an arbitrary delay d and up to an arbitrary constant c_d. From now on, a combined impulse response system of the form (1) will be referred to as a perfect equalization point (p.e.p.).

Consider a cost function $J \geq 0$ that is unimodal in the CSD. (By unimodal we mean that the function has a single minimum at a given p.e.p, and that the gradient of J in the CSD is zero at this point. Other points where the gradient is zero may exist, but correspond in this unimodal case to saddle points).

Let us consider the case when the equalizer and the channel impulse response both have finite length N and L respectively. Denote $\mathrm{W} = [w_0 \, \ldots \, w_{N-1}]^T$ and $\mathrm{H} = [h_0 \, \ldots \, h_{L-1}]^T$ the equalizer and channel impulse response coefficients respectively. It can be shown that in this case, the linear operator $\mathcal{H}$ takes the form of a rectangular matrix of size $N \times (N + L - 1)$ (with more columns than rows), which is not full column rank [6], thus has a right null space. Most criteria used in blind equalization can express their gradient with respect to (w.r.t.) the CSD in terms of their gradient w.r.t. the equalizer coefficients (EC):

$$\frac{\partial J}{\partial W^H} = \mathcal{H}^* \frac{\partial J}{\partial C^H} \tag{2}$$

Equation (2) implies that the gradient of the function in terms of the EC can be zero even if we do not have a null gradient in the CSD, i.e.: **1)** the gradient cancellation is not stable under the considered linear transformation, because of the null space of $\mathcal{H}$, and **2)** unimodality in the CSD does not ensure the existence of unimodality in the ED, which means that there is a possibility of having local minima in the ED even if the cost function is unimodal in the CSD.

Now consider a convex cost function $J \geq 0$ in the CSD and that its minimum is achieved for a given p.e.p. A useful property of convexity is that it is stable under the linear transform we are considering. Then, if $J(C)$ is convex in C, this implies that all minima in terms of the EC that could be introduced by the loss of rank of $\mathcal{H}$ correspond to the global one in terms of the combined system. As a consequence, all of them are p.e.p., and the problem of local minima does not exist anymore. Thus, under convexity of $J(C)$ the problem of misconvergence of the algorithms (which run in terms of the EC) towards a stable local minimum would be avoided. A first example of a convex cost function is found in [7].

To study function J in the ED, we have to express it in terms of W. Using the linear relationship $C = \mathcal{H}^T W$, we define $K(W) = J(\mathcal{H}^T W)$. Consider any two vectors W_1, W_2, and the corresponding combined systems $C_1 = \mathcal{H}^T W_1$, $C_2 = \mathcal{H}^T W_2$. If the convexity property holds for J in the CSD, for any value of $\forall\lambda \in (0, 1)$, we have $\lambda J(C_1) + (1-\lambda)J(C_2) \geq J(\lambda C_1 + (1-\lambda)C_2)$. And in terms of W, the same equation reads:

$$\begin{aligned}
&\lambda J(\mathcal{H}^T W_1) + (1-\lambda)J(\mathcal{H}^T W_2) \geq \\
&\quad J(\lambda\mathcal{H}^T W_1 + (1-\lambda)\mathcal{H}^T W_2) = J(\mathcal{H}^T(\lambda W_1 + (1-\lambda)W_2)) \\
&\quad \lambda K(W_1) + (1-\lambda)K(W_2) \geq K(\lambda W_1 + (1-\lambda)W_2)
\end{aligned}$$

As a result, convexity of $J(C)$ and the linear relation between W and C, also imply convexity of $K(W)$. The problem is that such type of convexity cannot hold in the whole CSD. A simple counterexample is that perfect equalization is defined up to an arbitrary delay. If a convex equalization criterion exists, any linear combination of two equalizers with different delays would also correspond to a p.e.p., and by recursion, any equalizer too. Clearly, this does not make sense. Thus, if we want to study convex cost functions, we have to analyze the situation for a fixed delay. In the rest of the paper, we consider convexity for a given equalization delay d. We choose to constrain the optimization in a non-convex subset of the CSD defined as $\mathcal{C}_d = \{C : |c_i| \leq |c_d|, |c_d| \neq 0\}$. The corresponding ED (also a non-convex subset) is $\mathcal{W}_d = \{W : \mathcal{H}^T\mathcal{W}_d \in \mathcal{C}_d\}$. The useful convexity results mentioned above for the ideal case obviously apply in these restricted cases, provided that one considers the right subset of the search space, as defined above.

Now consider the Shtrom-Fan (SF) cost functions with the above considerations in mind.

3. STUDY OF THE SF COST FUNCTION

The family of the SF cost functions [1] is based on the minimization of the distance between two norms:

$$J_F(C) \quad = \quad ||C||_p^\zeta - ||C||_q^\zeta \geq 0, \;\text{ for } 0 < p < q \qquad (3)$$

where the p norm is defined as $||C||_p^\zeta = \{\sum_n |c_n|^p\}^{\frac{\zeta}{p}}$, and p, q, ζ are any real number. It is shown [1] that the function reaches its minimum, i.e. $J_F(C) = 0$ in the noiseless case iff C is a p.e.p. of the form (1). Now, consider a particular case of this family of cost functions. In [8] it is claimed that:

$$J_F(C) = ||C||_1 - ||C||_\infty = \sum_k |c_k| - \max_k |c_k| \qquad (4)$$

is convex in the CSD, and in framework of this paper, it is easily seen that the criterion is, indeed, convex for a given equalization delay d (i.e. in the domain $\mathcal{W}_d$). However, practical use of the cost function (4) is difficult due to the infinite norm in the criterion. Hence, if one generalizes this case (the function is unimodal for a given delay), one could conclude that sectional convexity means having a function that is convex for a fixed delay in the CSD. We could suppose that the same definition applies to the rest of the class but we show below that this is not the case. (*Remark:* The problem about how a delay can be fixed practically in the CSD is addressed in section 3.3. For the moment, admit that it is fixed). From now on, concentrate on the case $p = 2, q = \zeta = 4$ (which is unimodal in the CSD [1]), a case of interest due to its similarities with the CMA [1], that will be analyzed in the CSD and then, see how the result translates in the ED.

3.1. The SF cost function for $p = 2, q = \zeta = 4$ in the CSD

The SF cost function for $p = 2, q = \zeta = 4$ is expressed as:

$$J_F(C) = \left(\sum_n |c_n|^2\right)^2 - \left(\sum_n |c_n|^4\right) \qquad (5)$$

Since we are studying an equalization function, we fix the equalization delay d, by requiring $|c_d| \geq |c_i|, \forall i \neq d$. Equation (5) can be rewritten as:

$$\begin{aligned}
J_F(C) &= \left(|c_d|^2 + \sum_{k\neq d}|c_k|^2\right)^2 - \left(|c_d|^4 + \sum_{k\neq d}|c_k|^4\right) \\
&= |c_d|^4\left(1 + 2\sum_{k\neq d}|\tilde{c}_k|^2 + \left(\sum_{k\neq d}|\tilde{c}_k|^2\right)^2\left(1 + \sum_{k\neq d}|\tilde{c}_k|^4\right)\right) \quad (6) \\
&= |c_d|^4 J_{NF}(\tilde{C}) \quad (7)
\end{aligned}$$

where $\tilde{C} = [\,\frac{c_1}{c_d}\,\ldots\,\frac{c_{d-1}}{c_d}\,\frac{c_{d+1}}{c_d}\,\ldots]$ (we omit $\tilde{c}_d = 1$ because it is constant). This leads to an implicit definition of the normalized cost function $J_{NF}(\tilde{C})$. We show below that $J_{NF}(\tilde{C})$ is convex for a given equalization delay, but that $J_F(C)$ is not, in general (in fact, counterexamples do exist, and one is shown below). Practically, this means that one has to fix also the value of the equalization gain c_d in order to obtain the desired convexity. Hence, we need to redefine the term sectional convexity.

Definition 1.1: For the general class of cost functions given by (3), the term sectional convexity in the CSD means to fix a delay d as well as the value of the coefficient associated to the delay d.

We also have to redefine the domains in which the search is undertaken. For a fixed delay d and a fixed value of $c_d \neq 0$, we shall be working in the convex subset $\mathcal{C}_{fd}$ of the CSD and its corresponding ED $\mathcal{W}_{fd}$ (also a convex subset of the ED) given by:

$$\begin{aligned}
\mathcal{C}_{fd} &= \{C : |c_i| \leq |c_d|\} \qquad (8) \\
\mathcal{W}_{fd} &= \{W : \mathcal{H}^T\mathcal{W}_{fd} \in \mathcal{C}_{fd}\} \qquad (9)
\end{aligned}$$

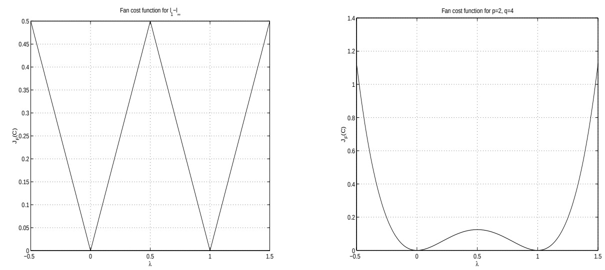
Obviously, the domains $\mathcal{C}_{fd}$ and $\mathcal{W}_{fd}$ are more restricted than the ones imposing only the delay. Using the new definition 1.1 and the redefined domains, it can be shown that:

Proposition 1: Using the definition 1.1, the SF cost function given by (7) is sectionally convex in $\mathcal{C}_{fd}$ given by (8) of the CSD.

Proof: See [9]. The convexity of (7), is obtained by proving the positivity of the hessian. An outline: derivate (7) twice with respect to all variables. The hessian is $H_{J_F} = 8\tilde{C}\tilde{C}^H +$

$4\left(1+\sum_m |\tilde{c}_m|^2\right) I - 12\,\mathrm{diag}|\tilde{C}|^2$, where I is the identity matrix and $|\tilde{C}|^2 = C \odot C^*$ where $\odot$ defines the Hadamar product. The positivity is proven in two steps: we split H_{J_F} in two matrices F_1, F_2 such that $H_{J_F} = 4(F_1 + F_2)$ as follows: $F_1 = \tilde{C}\tilde{C^H} + \sum_m |\tilde{c}_m|^2 I - 2\mathrm{diag}|\tilde{C}|^2, F_2 = \tilde{C}\tilde{C^H} + I - \mathrm{diag}|\tilde{C}|^2$. Since both of them can be shown to be positive, then the hessian H_{J_F} is also positive.

The positivity of H_{J_F} proves that (7) is convex in $\mathcal{C}_{fd}$. This implies (through the application of convexity) that all local minima that could exist are global ones over $\mathcal{C}_{fd}$ (the unique minimum is given by $\tilde{C} = 0$).



(a) $J_F(C) = ||C||_1 - ||C||_\infty$ (b) $J_F(C) = ||C||_2^4 - ||C||_4^4$

Fig. 2. **a)**(4) in the CSD between $C_1 = [1, 0, 0]$ and $C_2 = [0, 1, 0]$, plotted along the line $Y = \lambda J_F(C_1) + (1-\lambda) J_F(C_2)$. **b)**(5) in the CSD between the p.e.p. C_1 and C_2 (plotted as in case **a)**).

Illustration of results. Figure (2a) shows the cost function (4) in the CSD. It plots the evolution of the criterion when going from a p.e.p. to another one with different delay: $C_1 = [1\,0\,0]$ and $C_2 = [0\,1\,0]$. At the border, there is no change of convexity for each delay. This illustrates the claim that the SF criterion is convex for $p = 1, q = \infty$. Figure (2b) plots (5) under the same circumstances. At the border, the transition is not as sharp as for (2a), and there is a change of convexity in the basin of attraction of each delay. This supports our claim that defining sectional convexity based only on delay is not feasible. Finally, figure (3a) shows (7) in the CSD for (5) when we fix $c_d = 1$. The convexity is clearly seen, and shows the need for fixing the value of the equalization gain in order to obtain the convexity.

3.2. The SF cost function in the ED

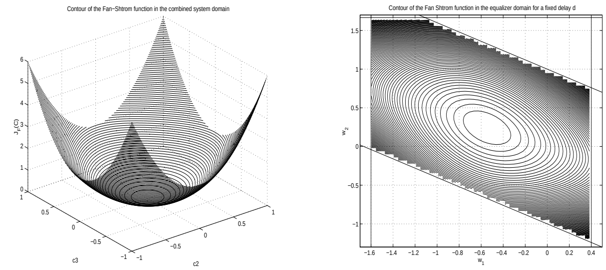
We now concentrate on how these results can be translated in the ED. In order to do so, we have to express (7) in terms of the equalizer. Using the results on convex functions (section 2) and the precise definition of sectional convexity, it is easily proven that:

Proposition 2: The SF cost function given by (7) is, by linearity, also convex in the ED given by the convex subset in (9).

An example of convexity in the ED is exhibited on figure (3b) The channel impulse response coefficients are $h = [1, 0.6]^T$ and we use an equalizer with three taps $W = [w_0\, w_1\, w_2]^T$. In this case, since we know the channel coefficients, it is very easy to control the domain $\mathcal{W}_{fd}$ such that $C = \mathcal{H}^T W$ belongs to $\mathcal{C}_{fd}$. In this case, we have fixed $c_0 = 1$, so that $w_0 = 1$. The cost function is plotted only in the domain defined by $\mathcal{W}_{fd}$ for clarity.

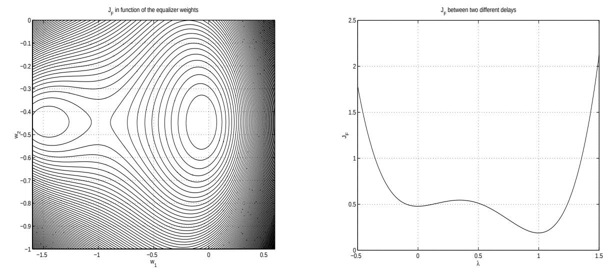
3.3. Comments about the obtained results

The practical use of the results is very difficult. In all cases considered, constraining the delay is required. When $p = 1, q = \infty$,



(a) Combined domain (b) Equalizer domain

Fig. 3. **a)** (7) for $d = 0, c_d = 1$ in the CSD given by $\tilde{\mathcal{C}}_{df} = \{C : |c_i| \leq |c_d|, c_d = 1\}$ for a vector C with 3 coefficients. **b)** (7) in region $\mathcal{W}_{df}$, for $d = 0, c_d = 1$. Domain $\mathcal{W}_{df}$ is defined for $w_0 = 1$. $\mathcal{W}_{df}$ does not cover all the values of $w_1,\ w_2$.



(a) (b)

Fig. 4. **a)**SF criterion (5) in the ED for W with $w_1 = 1$ and $h = [1, 0.6]^T$.**b)**Same function plotted between the two minima of fig. **a)**, which correspond to two different delays in the CSD.

this is enough to ensure convexity, while it is not the case when $p = 2, q = \zeta = 4$. But, putting a constraint on this delay is difficult: the idea used in [1] was to fix the middle coefficient of W to 1, a measure also used to avoid the trivial null solution of (5). But this only approximately ensures that a delay d will be fixed in the CSD. An example is shown on figure (4a), where we plot (5). We use $h = [1, 0.6]^T$ and an equalizer with three taps, the middle one being fixed to unity. We observe two minima, corresponding to different delays. Figure (4b) plots a view of the transition border between the two minima. This example shows that fixing a delay in the CSD is not ensured by simply fixing the value of an equalizer coefficient. A clean constraint should restrict the search in the region defined by $\mathcal{W}_d$, that requires the knowledge of the channel. Fixing the value of the equalization gain in the CSD, required for (5) is not easier: Once selected the value of d and restricted the optimization in the region (8), you must replace the d^{th} inequality constraints in (9) by an equality constraint, but for that, we require the knowledge of the channel, and this cannot be of practical use.

Since (5) is very similar to the CMA, we now study the CMA under the same conditions.

4. STUDY OF CMA COST FUNCTION IN THE CSD

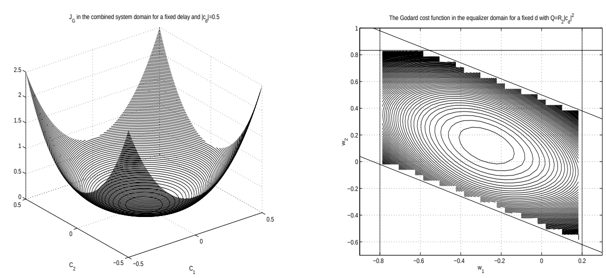
The Godard cost function for $p = 2$ [3] (or CMA $2-2$ [4]) is given by $J_G = E\{(|z_n|^2 - R_2)^2\}$ where z_n is the equalizer output and $R_2 \geq 0$ is a constant whose value is a function of the statistics of $\{a_n\}$ ($\{a_n\}$ is a zero mean i.i.d. symmetrical complex sequence with $E\{a_n^2\} = 0$). Let us rewrite J_G as follows:

$$J_G = E\{(|z_n|^2 - Q)^2\} \tag{10}$$

where Q is a constant. It can be easily proven that the gradient of (10) in the CSD, is null for a p.e.p. when $Q = \frac{E\{|a_n|^4\}}{E\{|a_n|^2\}}|c_d|^2$. Let $A = 2\left(E\{|a_n|^2\}\right)^2, B = E\{|a_n|^4\}$. Calculating the first derivative w.r.t. all the variables: $\frac{\partial J_G}{\partial c_i^*} = 4(B-A)|c_i|^2 c_i + 4Ac_i \sum_m |c_m|^2 - 4QE\{|a_n|^2\}c_i = 0$. Suppose $c_i \neq 0$. Solving w.r.t. a p.e.p. (with $c_i = c_d$) we find that the gradient is null when $Q = \frac{E\{|a_n|^4\}}{E\{|a_n|^2\}}|c_d|^2$. This was unnoticed in [3] which was written under the assumption that $|c_d|^2 = 1$. We note $Q = R_2|c_d|^2$. The following proposition holds.

Proposition 3: CMA $2-2$ is also sectionally convex, according to definition 1.1, in the CSD given by (8), and by linearity, also convex in the corresponding ED given by (9).

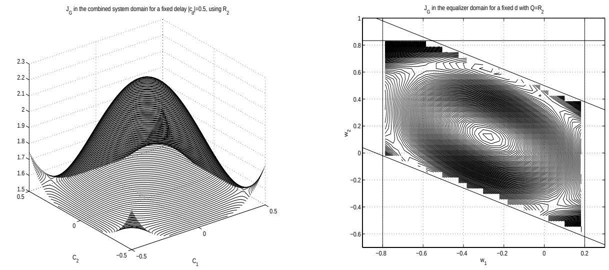
Proof: See [9]. An outline: the hessian of the CMA in the CSD is: $H_G = 8A\tilde{C}\tilde{C}^H + 12(B-A)\mathrm{diag}|\tilde{C}|^2 + 4(A - R_2E\{|a_n|^2\})I + 4A\gamma I$, where $\gamma = \sum_m |\tilde{c}_m|^2$. Dividing by $4A$ and substituting the value of R_2: $H_G^{'} = 2CC^H + 3(K-1)\mathrm{diag}|C|^2 + (1-K)\,I + \gamma I$, where $K = \frac{B}{A}$. K represents the value of the kurtosis of the input signal. Many constellations (such as PSK and QAM) are subgaussian, i.e. $\frac{1}{2} \leq K \leq 1$. The positivity is proven in two steps: we split H_G in two matrices G_1, G_2 such that $H_{J_G} = G_1 + G_2$ where: $G_1 = \tilde{C}\tilde{C}^H + \gamma I - 2\,\mathrm{diag}|\tilde{C}|^2, G_2 = \tilde{C}\tilde{C}^H + (3K-1)\mathrm{diag}|\tilde{C}|^2 + (1-K)\,I$. $G_1 = F_1$, which is positive, and a sufficient condition for G_2 to be positive is that $\frac{1}{2} \leq K \leq 1$.



(a) Combined Domain (b) Equalizer Domain

Fig. 5. **a)**CMA in the CSD for a vector C with 3 coefficients and $c_d = .5, Q = R_2|c_d|^2$. **b)** CMA in the corresponding ED $\mathcal{W}_{df} = \{W : \mathcal{H}^T W \in \mathcal{C}_{df}\}$, for $d = 0, c_d = .5, Q = R_2|c_d|^2$. Convexity is clearly seen in both domains

Illustration of results. We used a QAM-4 signal, so $R_2 = 2$. Figures (5a) and (5b) show the sectional convexity of J_G in both the CSD and the ED when $c_d = 0.5$ and $Q = R_2|c_d|^2 = 0.5$. The case $c_d = 1$ is not shown since is the same as the SF case shown in figure (3a). The convexity in both domains is clearly seen. Hence, we have here the same conclusion as for criterion (5): under subgaussian assumption, the CMA $2-2$ is sectionally convex in the CSD (8), and by linearity in the corresponding ED (9), provided that one is able to constrain the value of c_d and Q. Notice the effects in the CSD and in the ED (figures (6a) and (6b)), when $c_d = 0.5$, but the value of Q is fixed incorrectly, as for example, $Q = R_2$. The adaptive algorithm may run, but since the convexity is lost in both domains, we will not be able to equalize and may converge anywhere. This shows that in order to be convex, we must fix the value c_d and Q correctly.



(a) Combined Domain (b) Equalizer Domain

Fig. 6. **a)** CMA in the CSD and in the corresponding ED **b)** when $c_d = 0.5$ but $Q = R_2$. The criterion is not convex anymore.

5. CONCLUSIONS

We have analyzed the SF cost function for a particular case, which is very similar to the CMA criterion, and found that is sectionally convex in the CSD, with a precise definition of sectional convexity. We also showed that these results can be translated by linearity to the corresponding ED, if we are able to fix the delay as well as the coefficient associated to the delay in the CSD. However, the result does not seem to be applicable unless we have access to the coefficients of the CSD or the channel impulse response ones to be able to impose the required constraints. Furthermore, this property was shown to be true also for the CMA criterion.

6. REFERENCES

[1] V. Shtrom and H. Fan, "New Class of Zero-Forcing Cost Functions in Blind Equalization," *IEEE Trans. on SP*, vol. 46, no. 10, pp. 2674–2683, Oct. 1998.

[2] Y. Sato, "A Method of Self-Recovering Equalization for Multilevel Amplitude-Modulation Systems," *IEEE Tr. on Com.*, vol. AC-25, no. 3, pp. 679–682, June 1975.

[3] D. Godard, "Self-Recovering Equalization and Carrier Tracking in Two-Dimensional Data Communications Systems," *IEEE Tr. on Com.*, vol. 28, no. 11, pp. 1867–1875, Nov. 1980.

[4] J.R. Treichler and B.G. Agee, "A New Approach to Multipath Correction of Constant Modulus Signals," *IEEE Trans. on SP*, vol. ASSP-31, pp. 349–472, April 1983.

[5] Z. Ding, R. A. Kennedy, B. D. O. Anderson, and C. R. Johnson Jr., "Ill-Convergence of Godard Blind Equalizers in Data Communications Systems," *IEEE Tr. on Com.*, vol. 39, no. 9, Sept. 1991.

[6] Z. Ding, C. R. Johnson Jr., and R. A. Kennedy, "On the (Non)Existence of Undesirable Equilibria of Godard Blind Equalizers," *IEEE Trans. on SP*, vol. 40, no. 10, Oct. 1992.

[7] S.Vembu, S. Verdu, R. Kennedy, and W. Sethares, "Convex cost functions in blind equalization," *IEEE Trans. on SP*, vol. 42, no. 8, Aug. 1994.

[8] V. Shtrom and H. Fan, "A Refined Class of Cost Functions in Blind Equalizations," *Proc. ICASSP*, pp. 2273–2276, 1997.

[9] M. Corlay, M.Charbit, and P. Duhamel, "Comments about New Class of Zero-Forcing Cost Functions in Blind Equalization," *to be submitted to IEEE Trans. on SP*.