New Issues on the Criterion Properties and Algorithm Convergence of the Generalized CM Approach

A. O. Neves*, M. D. Miranda**, J. M. T. Romano*

*DSPCom, Faculty of Electrical and Computer Engineering, State University of Campinas, SP, Brazil, P.O. 6101, CEP 13083-970 Phone: +55-19-37883703, Fax: +55-19-32891395
** Universidade Presbiteriana Mackenzie, São Paulo, SP, Brazil, CEP 01302–907, Phone: +55-11-2368711
Fax: +55-11-2368600
aline@decom.fee.unicamp.br, mdm@mackenzie.com.br, romano@decom.fee.unicamp.br

Abstract- **The first aim of this work is to extend some recent and important results on the equivalence between Shalvi-Weinstein and Constant Modulus methods, by including the so called Generalized Constant Modulus criterion and its corresponding stochastic algorithm. Such method avoid the power restriction on the optimization process, based on the idea that any constant modulus signal at the output of the equalizer represents the achievement of equalization. We show the equivalence between this alternative criterion and the Shalvi-Weinstein one. As far as the corresponding algorithm is concerned, we introduce a study on an alternative stochastic approximation, which leads to an increased performance in terms of convergence rate. The behavior is analyzed in terms of the stochastic gradient vector and evaluated as a function of the algorithm parameters. The dependence of the algorithm in terms of its initial conditions is also analyzed and a modified center-spike procedure is proposed to attain an improved convergence behavior.**

1. INTRODUCTION

The interest in blind deconvolution techniques for digital communications systems has considerably increased in the last years, as well as the proposition of new algorithms and analysis tools. For that reason, it is important to investigate the similarities and differences between these new approaches and some already consolidated techniques. This is the main motivation of the present paper.

For instance, as far as blind deconvolution criteria are concerned, the relationships between Shalvi-Weinstein [5] and Godard [2] criteria have been stated in [6]. Such equivalence takes into account the power restriction, which poses that the equalizer output and the transmitted signal have the same power. A first issue of this paper is to extend the study in [6] to the case where the power restriction is not considered. With this objective, we consider the method proposed in reference [1], which leads to an alternative CMA-based algorithm. Then we show that the criterion used in [1] is equivalent to the Shalvi-Weinstein criterion, without the power restriction.

As far as adaptive algorithms are concerned, references [8] and [9] show that super-exponential and gradient search methods are equivalent. Clearly CMA [2,3] is included in such analysis since it is a gradient search algorithm. Then, in the light of these works, we study the case of the algorithm in [1], where the constant modulus condition is generalized by the absence of power restriction. We show that the interesting performance of this technique is mainly due to a non-conventional stochastic approximation method, employed in [1] with no major analysis.

Such study allowed us to provide a more systematic comparison between the algorithm in [1] and the classical CMA and Super-Exponential (SEA) algorithms [4]. It also opens perspectives in obtaining effective algorithms, in terms of convergence rate, by using non-conventional stochastic approximation methods.

This paper is organized as follows. Section 2 presents a brief review on the relationships between the Constant Modulus (CM) and the Shalvi-Weinstein (SW) criteria. The generalized constant modulus (GCM) scheme proposed in [1] is revisited in Section 3. Section 4 discusses the non-conventional stochastic approximation in GCM algorithm. The performance of such algorithm is evaluated and compared with the CM and SW techniques in Section 5. Section 6 discusses the effect of initialization on the algorithm as well as a modification in the center-spike procedure, which results in an improved convergence rate. Finally the conclusions are presented in Section 7.

2. A BRIEF RECALL ON SHALVI-WEINSTEIN AND CONSTANT MODULUS CRITERIA

The constant modulus criterion is given by [3]:

$$J_{CM} = E\left[\left(\left|y_n\right|^2 - r\right)^2\right], \tag{1}$$

where $E[.]$ is the expectation operator, y_n is the equalizer output signal and r is given by:

$$r = E\left[\,|\,a_n\,|^4\right]\,/\,E\left[\,|\,a_n\,|^2\right] \tag{2}$$

where a_n is the transmitted symbol.

This criterion penalizes deviations of the equalizer output signal from a constant modulus defined by r.

The Shalvi-Weinstein criterion is given by [4]:

$$J_{SW} = \frac{C_4^y}{\left(C_2^y\right)^2} = \frac{E\left[\left|y_n\right|^4\right]}{E^2\left[\left|y_n\right|^2\right]} - k, \tag{3}$$

where $C_2^y = E\left[|y_n|^2\right]$, $C_4^y = E\left[|y_n|^4\right] - kE^2\left[|y_n|^2\right]$ is the fourth order cumulant of the equalizer output signal, k=3 when y_n is a real sequence and k=2 when the process is complex and circularly symmetric.

Now let us consider that $E\left[|y_n|^2\right] = E\left[|a_n|^2\right]$ holds for SW criterion. It means that the equalizer output has the same power of the transmitted signal. In this case, it is shown in [6] that CM and SW criteria have the same convergent points. Thus the resultant algorithms search for the optimal points of the same cost function, which belongs to the family of Donoho's deconvolution objective function [7].

The CM [2,3] and the SW [4] algorithms can be respectively written as:

$$\mathbf{w}_n = \mathbf{w}_{n-1} - \mu y_n \mathbf{u}_n^* \left(|y_n|^2 - r\right) \qquad (4)$$

and

$$\mathbf{w}_n = \mathbf{w}_{n-1} + \delta \mathbf{Q}_n \mathbf{u}_n^* y_n \left(|y_n|^2 - r\right) \qquad (5)$$

where $y_n = \mathbf{w}_n^T \mathbf{u}_n$ is the equalizer output signal, $\mathbf{w}_n=[w_0 \ w_1 \ ...w_{L-1}]^T$ is the equalizer tap vector, $\mathbf{u}_n=[u_n \ u_{n-1} \ ...u_{n-L+1}]^T$ is the equalizer input signal vector, μ is the adaptation step size, $\delta = C_2^a / C_4^a$ and $\mathbf{Q}_n$ estimates the inverse of the autocorrelation matrix of the equalizer input, updated by:

$$\mathbf{Q}_n = \frac{1}{1-\beta}\left(\mathbf{Q}_{n-1} - \frac{\beta \mathbf{Q}_{n-1}\mathbf{u}_n^* \mathbf{u}_n^T \mathbf{Q}_{n-1}}{1-\beta+\beta \mathbf{u}_n^T \mathbf{Q}_{n-1}\mathbf{u}_n^*}\right) \qquad (6)$$

where β is constant. The role of matrix $\mathbf{Q}_n$ is to provide a prewhitening operation over the channel output [5]. It is worth to note that the gradients of (1) and (3) are respectively given by:

$$\frac{\partial J_{CM}}{\partial \mathbf{w}} = E\left[|y_n|^2 y_n \mathbf{u}_n^*\right] - r\, E\left[y_n \mathbf{u}_n^*\right] \qquad (7)$$

and

$$\frac{\partial J_{SW}}{\partial \mathbf{w}} = \frac{1}{E^2\left[|y_n|^2\right]}\left(E\left[|y_n|^2 y_n \mathbf{u}_n^*\right] - r^y E\left[y_n \mathbf{u}_n^*\right]\right) \qquad (8)$$

where $r^y = E\left[|y_n|^4\right] / E\left[|y_n|^2\right]$.

Then it can be noted that the main difference between expressions (7) and (8) is the term r^y in (8) instead of the constant r in (7). However, when equalization is achieved, the following equality holds:

$$r^y = \frac{E\left[|y_n|^4\right]}{E\left[|y_n|^2\right]} = \frac{E\left[|a_n|^4\right]}{E\left[|a_n|^2\right]} = r \qquad (9)$$

In such conditions, it comes from equations (7) and (8) that the gradients of both criteria CM and SW will vanish for the same stationary points.

To replace the term r^y by r corresponds to previously pose the power restriction in the SW gradient computation. That is effectively done when SEA is derived from the SW criterion [4]. Thus, we can confirm the result in [6] that CMA and SEA with the power restriction will converge to the same stationary points, since both adaptation process deal with equivalent gradient vectors.

3. GENERALIZING THE CONSTANT MODULUS

As well as replacing r^y by r in (8) leads to impose the power restriction on the SW criterion, we could untie the fixed constant modulus r in (1), so that an alternative algorithm could be carried out without power restriction. This could be achieved by replacing r by a term depending on y. However, the capacity of recovering the constant modulus property at the equalizer output must be preserved.

In this sense reference [1] proposed the following equalization criterion:

$$J_{GCM} = \frac{E\left[\left(|y_n|^2 - E\left[|y_n|^2\right]\right)^2\right]}{\left(E\left[|y_n|^2\right]\right)^2}, \qquad (10)$$

so-called the Generalized Constant Modulus (GCM) criterion, because it did not establish a fixed value for the equalizer output modulus. Its development was based on the idea that any constant modulus signal, at the output of the equalizer, represents the achievement of equalization. A simple automatic gain control (AGC) can then correct the output signal level, if necessary.

The GCM approach has been successfully applied in different cases, including space-time processing [10]. Nevertheless a more theoretical investigation about the corresponding criterion and adaptive algorithm, in the light of the classical techniques, had not yet been carried out.

As far as the GCM criterion is concerned, it is interesting to compare (10) with the SW criterion in (3), since both expressions do not consider the power restriction. After some manipulations it comes:

$$J_{GCM} = \frac{E\left[|y_n|^4\right]}{E^2\left[|y_n|^2\right]} - 1 = J_{SW} + k - 1 \qquad (11)$$

Thus, a clear equivalence holds, so that GCM criterion seeks the same optimal points of SW criteria, what includes it into the family of Donoho's deconvolution objective function.

Hence, the algorithm resulted from GCM criterion can be considered as belonging to the Shalvi-Weinstein family of algorithms, but without the power restriction. As shown in the next section, such algorithm is derived from an instantaneous version of (10), so that a most careful study about this alternative stochastic approximation must be carried out in order to well evaluate the algorithm behavior.

4. THE GCM STOCHASTIC ALGORITHM

As posed in [1], the criterion in (10) is related to the signal-to-envelope variation ratio (SVR), a parameter for signal quality measurement given by:

$$SVR\alpha \frac{E^2\left[|y_n|^2\right]}{\left(|y(n)|^2 - E\left[|y_n|^2\right]\right)^2} \quad (12)$$

The SVR expresses how close a signal is from a constant modulus property.

Based on the fact that SVR is an instantaneous measure, the so-called Generalized Constant Modulus Algorithm (GCMA) was derived in [1] by posing the stochastic approximation over the original criterion in (10) and not over its corresponding gradient. For a number of cost functions, both procedures lead to the same result. This is not the case of the GCM function and such feature has shown to be rather favorable as far as the algorithm performance is concerned, as shown in the sequel.

First let us define the following stochastic criterion, from (10):

$$J'_{GCM} = \frac{\left(|y_n|^2 - E\left[|y_n|^2\right]\right)^2}{\left(E\left[|y_n|^2\right]\right)^2}, \quad (13)$$

what leads to:

$$\frac{\partial J'_{GCM}}{\partial \mathbf{w}} = \frac{1}{\bar{z}^2}\left(|y_n|^2 y_n \mathbf{u}_n^* - \mathbf{b}\hat{r}^y + |y_n|^2 \mathbf{b} - y_n \mathbf{u}_n^* \bar{z}\right), \quad (14)$$

where $\bar{z} = E\left[|y_n|^2\right]$, $\mathbf{b} = E\left[y_n \mathbf{u}_n^*\right]$ and $\hat{r}^y = |y_n|^4 / E\left[|y_n|^2\right]$.

Then GCMA is given by:

$$\mathbf{w}_n = \mathbf{w}_{n-1} - \mu \frac{\partial J'_{GCM}}{\partial \mathbf{w}}, \quad (15)$$

where μ is the adaptation step size. The terms $\bar{z}$ and $\mathbf{b}$ are calculated by time averages:

$$\begin{aligned} \bar{z}_n &= \lambda \bar{z}_{n-1} + (1-\lambda)|y_n|^2 \\ \mathbf{b}_n &= \lambda \mathbf{b}_{n-1} + (1-\lambda) y_n \mathbf{u}_n^* \end{aligned} \quad (16)$$

where λ is the forgetting factor.

To provide a performance evaluation of GCMA and compare this algorithm with CMA and SEA, it is interesting to access the behavior of the corresponding gradient vector. It is easily seen that equation (14) can be separated in two terms:

$$\left(\underbrace{|y_n|^2 y_n \mathbf{u}_n^* - \hat{r}^y E[y_n \mathbf{u}_n^*]}_{TI} + \underbrace{|y_n|^2 E[y_n \mathbf{u}_n^*] - y_n \mathbf{u}_n^* E\left[|y_n|^2\right]}_{TII}\right) \quad (17)$$

The first one, *TI*, is similar to (7) and (8), while the second one, *TII*, does not appear in those equations. It can be shown that *E[TI]* corresponds to the SW gradient vector while *E[TII] = 0*. This indicates that GCMA is expected to have a different behavior during the transient period. Meanwhile the effect of vector *TII* is not so evident and must be better evaluated by simulations.

5. PERFORMANCE EVALUATION AND SIMULATION RESULTS

In the presented simulations, the following measure of intersymbol-interference (ISI) was used to evaluate the equalization performance:

$$\text{ISI}(s) = \frac{\sum_n |s_n|^2 - |s_{\max}|^2}{|s_{\max}|^2}, \quad (18)$$

where s_n is the n[th] element of the impulse response of combined (channel + equalizer) system and $s_{\max}$ is the maximal absolute value of s_n.

Taking the value of ISI achieved by the SEA as a reference, and posing all adaptation parameters so that the algorithms converge to the same ISI, to allow a correct comparison of their convergence speed, we obtained the results in Figures 1-3. Such figures correspond each one to a distinct non-minimum phase channel, for which the zeros are respectively given by 0.2 and 5; 0.5 and 2; and 0.8 and 1.25. To simplify the correct choice of the adaptation step sizes, the equalizer input was adjusted to have unit power. The equalizer was kept with 15 taps and the modulation of the transmitted symbols was BPSK. All algorithms were initialized using the *center-spike* method. Furthermore, there was no addition of noise.

In figure 1 the channel zeros are closer to the origin of the z-plane, what corresponds to a more uncorrelated signal at the equalizer input. In such case the three algorithms have similar performances and CMA and GCMA present almost overlapped curves.

As the channels zeros move away from the origin (figure 2), the performance of GCMA becomes clearly superior if compared to CMA.

Finally, for zeros which are still closer to the unity circle (figure 3), what represents the case of most correlated channel output signal, the superiority of SEA comparing to the two others is significant. Moreover, there is a tendency in coming back to a situation of similar performance between GCMA and CMA.

When considering additive white gaussian noise, we obtained basically the same results. SEA remains the faster algorithm, followed by GCMA and CMA presents the slowest convergence. Once again, GCMA and CMA performances tend to approximate as the channels zeros get closer to the origin or to the unit circle. As an example, figure 4 shows the result obtained using a channel with zeros in 0.8 and 1.25, a signal to noise ratio (SNR) of 20dB and a 15 taps equalizer.

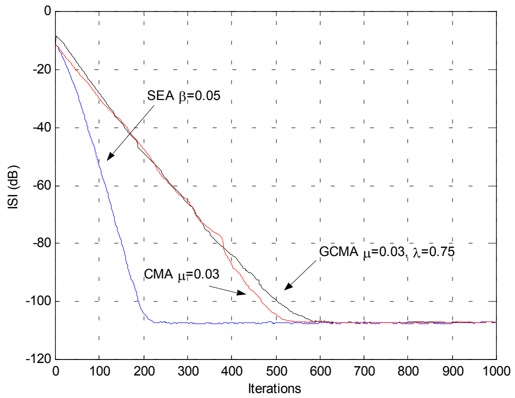

Figure 1: Performance of the algorithms for a channel with impulse response h(z)=0.1856-0.9650z^{-1}+0.1856z^{-2}, no noise

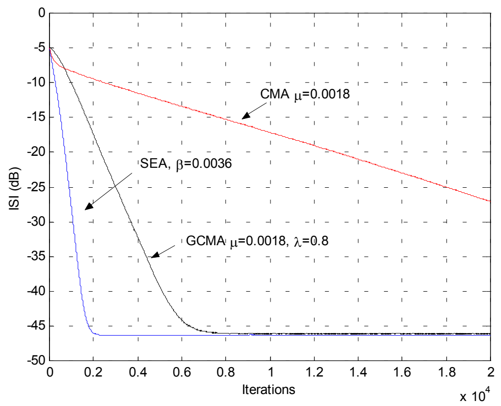

Figure 2: Performance of the algorithms for a channel with impulse response h(z)=0.3482-0.8704z^{-1}+0.3482z^{-2}, no noise

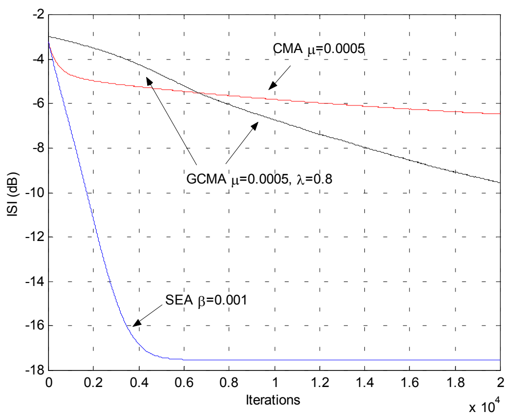

Figure 3: Performance of the algorithms for a channel with impulse response h(z)=0.4015-0.8231z^{-1}+0.4015z^{-2}, no noise

Hence, as far as the convergence rate is concerned, it is interesting to point out two conclusions from the set of presented results:

i- SEA presents the better performance, converging faster and such difference is more significant as the zeros of the channel are closer to the unity circle

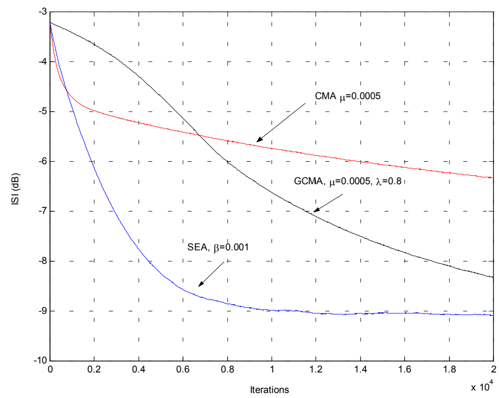

Figure 4: Performance of the algorithms for a channel with impulse response h(z)=0.4015-0.8231z^{-1}+0.4015z^{-2}, SNR=20dB

ii- GCMA presents an improved convergence rate when compared to CMA, except for zeros too close to the origin or to the unity circle.

In fact the three algorithms are gradient search approaches. However SEA makes use of matrix $\mathbf{Q_n}$ to provide a prewhitening operation, which improves the convergence rate with an increasing computational complexity.

Now, to confirm the effect of vector *TII* in the convergence rate of GCMA, we simulated a reformulated version of the algorithm in (15), removing the term *TII*. Such algorithm, named GCMA$_I$ is given by:

$$\mathbf{w}_n = \mathbf{w}_{n-1} - \frac{\mu}{\bar{z}^2}\left(\left|y_n\right|^2 y_n \mathbf{u}_n^* - \mathbf{b}\hat{r}^y\right) \qquad (19)$$

The result is shown in figure 5. The system configuration was the same as the one used to obtain figures 1-3 and the channel used had the following impulse response: h(z)=0.3482-0.8704z^{-1}+0.3482z^{-2}. It can be seen that the algorithm had a very poor performance.

Hence, the term *TII* provides like a favorable preprocessing over the equalizer input signal. However such effect becomes negligible for uncorrelated or too correlated channel outputs.

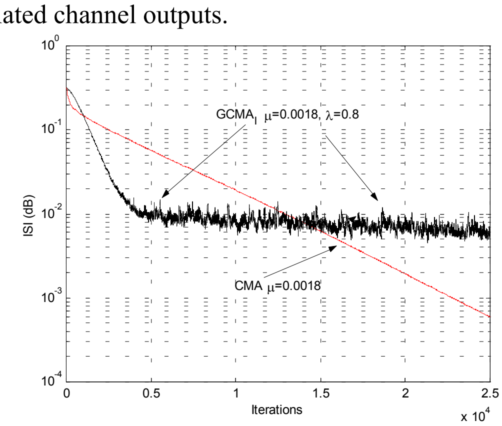

Figure 5: Comparing the performance of GCMA$_I$ and CMA, h(z)=0.3482-0.8704z^{-1}+0.3482z^{-2}, no noise

It is worth pointing out that the term *TII* comes from the alternative stochastic approximation method, which is the main feature in the GCMA derivation. As far as computational complexity is concerned, GCMA cost is similar to CMA and rather lower than the SEA one.

6. GCMA Initialization

The effect of initialization in GCMA algorithms must be investigated in the light of its corresponding criterion, the characteristics of which were discussed in section 3.

In fact, due to the non-existence of the power restriction, GCM and SW criteria have an infinite number of possible solutions. Taking into account the relationship between CM and SW with the power restriction, studied in [6], it is expected that the CM minima will be one of the GCM possible solutions. Figure 6 illustrates such consideration. We plot the GCMA convergence paths for a 2 taps equalizer and a channel given by $h(z)=1-0.6z^{-1}$. The CM global and local minima are also indicated in the same figure.

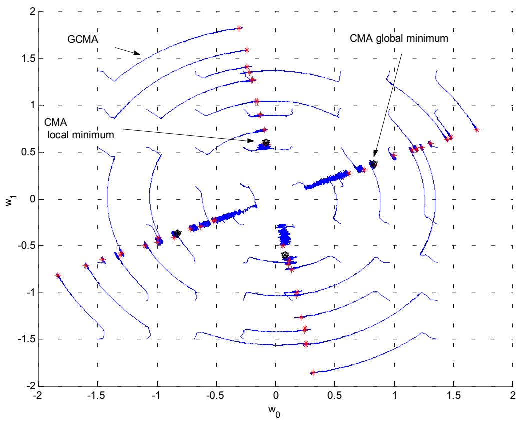

Figure 6: The convergence of GCM algorithm, μ=0.0025, λ=0.97

We can see that the GCMA convergence point depends on the initialization and may assume a value belonging to a valley of solutions. This valley includes the solutions obtained by CM criterion.

In section 5, all simulations were carried out using the *center spike* initialization, which consists of setting all taps to zero but one, that assumes unity value. However, we can expect that, by initializing GCMA closer to its valley of solutions, the algorithm will converge faster. The same would happen if we initialized CMA near its global minimum. The drawback in this last case is that the solution is not known. On the other hand, from figure 6, we observe that the two GCMA valleys of solutions cross the origin of the plane defined by the equalizer taps. Then if initial taps are set closer to the origin, they will also be closer to the valley of solutions.

To confirm such affirmation, we initialized the algorithm with *center spike*, but setting the non-zero tap to 0.3 instead of 1. The results are shown in figures 7 and 8, respectively for non-minimum phase channels having zeros in 0.2 and 5; 0.5 and 2. The equalizer had 15 taps, no noise was considered and the transmitted signal modulation was BPSK.

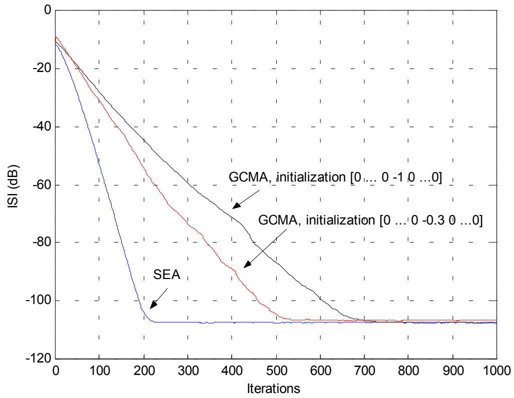

Figure 7: Changing the initialization of GCMA; $h(z)=0.1856-0.9650z^{-1}+0.1856z^{-2}$, GCMA: μ=0.0035, λ=0.7; SEA: β=0.06, no noise

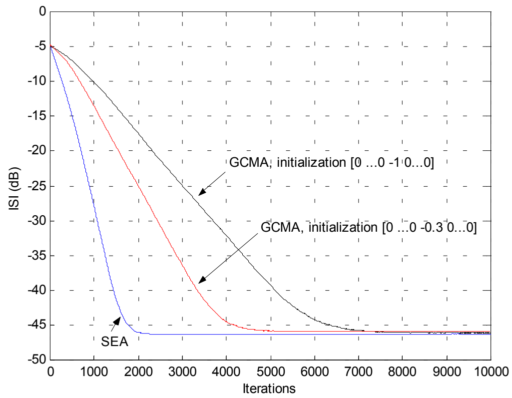

Figure 8: Changing the initialization of GCMA; $h(z)=0.3482-0.8704z^{-1}+0.3482z^{-2}$, GCMA: μ=0.00025, λ=0.8; SEA: β=0.0036, no noise

Figures 7 and 8 indicates a real gain in performance, particularly in figure 8, where the algorithm with the new initialization converges 2500 iterations before the traditional one. Clearly GCMA still does not achieve the SEA performance, but it becomes even better than CMA.

However, it is important to note that when all equalizer taps are set to zero, the denominator of GCMA cost function in (10) vanishes. For that reason, the non-zero tap at the initialization procedure should not have a very small value. As this value approaches zero, the algorithm becomes more unstable, requiring a smaller value for the adaptation step size μ. By simulations, we verified that for values around 0.3 a good trade-off between convergence rate and stability could be attained.

7. Conclusion

The interest in investigating the similarities and differences between new and consolidated techniques of blind equalization is the main motivation of the present work. In this sense, a theoretical study and a performance evaluation of the so-called GCM criterion and algorithm was carried out in the light of the Shalvi-Weinstein criterion and some previous results on SW and CM equivalencies.

The GCM criterion was shown to be equivalent to the SW one and its corresponding algorithm (GCMA) was more deeply studied. The new obtained results concern its non-usual stochastic approximation and the dependence on initialization. As a consequence, a modified center-spike procedure was indicated with a significant gain in the convergence rate.

Comparing it with CMA and the SEA, we concluded that GCMA usually outperforms CMA, due to the effect of the alternative stochastic approximation method. The interest is that such performance is achieved without using a prewhitening auto-correlation matrix, i.e., with a lower computational cost when compared to the SEA.

The overall results of this paper confirm the interest in dealing with a more unified view on the different blind equalization methods, in order to improve the theoretical framework of this field as well as to indicate alternative and efficient new techniques.

Acknowledgments

We would like to manifest our appreciation to the Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP) for supporting this work.

References

[1] F. R. Cavalcanti, A.Brandão, J. M. Romano, *A Generalized Constant Modulus Algorithm for Blind Equalization*, Global Telecommunication Conference (GLOBECOM), 1998

[2] D. Godard, *Self-Recovering Equalization and Carrier Tracking in Two-Dimensional Data Communication Systems*, IEEE Trans. on Communications, Vol. COM-28, Issue 11, Nov. 1980, Pages 1867-1875

[3] J. Treichler.; B. Agee, *New Approach to multipath correction of constant modulus signals*, IEEE Trans. Acoust., Speech and Signal Processing, Vol ASSP-31, No 2, April, 1983, pp. 459-472

[4] O. Shalvi; E. Weinstein, *Super-Exponential Methods for Blind Deconvolution*, IEEE Trans. on Information Theory, Volume 39, Issue 2, March 1993, Pages 504-519

[5] O. Shalvi; E. Weinstein, *New Criteria for Blind Deconvolution of Nonminimum Phase Systems (Channels)*, IEEE Trans. on Information Theory, Vol. 36, Issue 2, March 1990, Pages 312-321

[6] P. Regalia, *On the Equivalence Between the Godard and Shalvi Weinstein Schemes of Blind Equalization*, Signal Processing, Volume 73, Issues 1-2, 2 January 1999, Pages 185-190

[7] D. Donoho, *On Minimum Entropy Deconvolution*, *in* Applied Time Series Analysis II, D. F. Findley, Ed. New York: Academic, 1981

[8] M. Mboup, P.A. Regalia, *A gradient search interpretation of the super-exponential algorithm*, IEEE Trans. on Information Theory, vol 46, No 7, pp. 2731-2764, Nov 2000

[9] P. A. Regalia, M. Mboup, *Undermodeled Equalization: A Characterization of Stationary Points for a Family of Blind Criteria,* IEEE Trans. on Signal Processing, Vol. 47, Issue 3, March 1999, pp. 760-770.

[10] F. R. Cavalcanti, A. L. Brandão, J. M. Romano, *The Generalized Constant Modulus Algorithm Applied to Multiuser Space-time Equalization*, Proc. 2nd. SPAWC, 1999.