Blind Deconvolution of Correlated Sources Based on Second-Order Statistics

Denis G. Fantinato, Romis Attux, Aline Neves, R. Suyama and J.M.T. Romano

Abstract—**The blind deconvolution of signals composed of statistically dependent samples is an important practical problem whose understanding still requires the clarification of many theoretical points. In this work, we present an analysis of this problem that includes two well-established methods - the canonical constant modulus algorithm (CMA) and a correntropy-based method - and two novel strategies that explore the temporal profile of the signal of interest. These techniques are compared in a number of representative scenarios, where it will be possible to form a clearer view of their potentialities and also of some peculiarities of the problem itself.**

Keywords—**Blind deconvolution, correlated sources, correntropy, constant modulus criterion.**

I. INTRODUCTION

The task of blind deconvolution (or equalization) has received a great deal of attention in the last decades [1]-[4]. From the perspective of digital communications, this interest can be justified by the fact that the operation of blind equalizers does not depend on the use of a reference signal, which is beneficial in terms of the achievable transmission rate.

Several methods and fundamental theoretical results helped establish the field of unsupervised signal processing [1], in which we highlight the constant modulus (CM) criterion [2]. The CM approach has been the object of several investigations in the context of a transmitted signal composed of mutually independent sources, but, in spite of efforts like [3][4], its behavior for non-independent sources still demands significant clarification. It should be mentioned that sources of the latter kind are practically important in view of the potential application of different types of codes before signal transmission or of the use of the CMA for handling analog discrete-time signal processing (e.g. in audio-related scenarios).

Having in mind the peculiarities of blind deconvolution of non-independent signals, in [5], a method was proposed that takes into account the time structure of the source of interest. This is achieved by making use of a metric belonging to the field of information theoretic learning (ITL): the correntropy. As the name indicates, the metric is an attempt to generalize the concept of correlation towards the inclusion of a richer palette of statistical properties of the signal of interest.

The use of correntropy in blind deconvolution was a significant step in the direction of a better understanding of the problem under non-independent signals, but there remain plenty of unanswered questions regarding the behavior of classical approaches within this formulation. Some of these questions were addressed in [6], which presents elements of comparison between the CM algorithm (CMA) and the correntropy-based method. In this work, however, a further step is taken with the aid of two novel criteria that, aside from their intrinsic relevance, serve as bridges to relate the CMA to the ITL solution. These are a modification of the Multiuser CMA (MU-CMA) [7] and a criterion based on the restoration of the source correlation profile. The associated algorithms will serve as comparative performance tools, which will bring to the fore the crucial trade-off between reachable performance and complexity.

This paper is organized as follows. A brief definition of the CM criterion and the correntropy-based method [5] are presented in Section II. In Section III, we present the modified version of the MU-CMA, followed by the definition of the source correlation profile restoration criterion in Section IV. The connections between the presented algorithms are explored in Section V. For a better understanding, an error surface analysis for a correlated and an uncorrelated case is shown in Section VI. In Section VII, a performance analysis is conducted for two representative scenarios. Finally, the conclusions and the possibilities for further work are summarized in Section VIII.

II. BLIND EQUALIZATION CRITERIA

In this section, we will discuss in more detail the two criteria that will serve as examples of blind strategies: the CM criterion and the correntropy-based criterion [5].

A. Constant Modulus Criterion

The CM criterion is based on the idea of penalizing deviations of a quadratic version of $y(n)$ from a fixed constant R_2. It can be defined by means of the minimization of a cost function, as follows:

$$\min_{\mathbf{w}} J_{CM}(\mathbf{w})) = \min_{\mathbf{w}} E\left[\left(|y(n)|^2 - R_2\right)^2\right], \quad (1)$$

where $E[\cdot]$ denotes statistical expectation, $\mathbf{w}$ is the finite impulse response (FIR) equalizer coefficient vector, $y(n)$ is the equalizer output signal, $R_2 = E\left[|s(n)|^4\right]/E\left[|s(n)|^2\right]$ and $s(n)$ is the transmitted signal. It is important to remark that this proposal was made under the assumption that the transmitted signal is composed of independent and identically distributed (i.i.d.) samples, i.e., non-dependent sources. When dealing with dependent data, previous efforts [3][4] revealed that, in

Denis G. Fantinato, Romis Attux and J.M.T. Romano, School of Electrical and Computer Engineering, University of Campinas, Campinas-SP, Brazil, E-mails: {denisgf,attux}@dca.fee.unicamp.br, romano@dmo.fee.unicamp.br. Aline Neves and R. Suyama, Engineering, Modeling and Applied Social Science Center, Federal University of ABC, Santo André - SP, Brazil, E-mails: {aline.neves,ricardo.suyama}@ufabc.edu.br This work was partially supported by CAPES and CNPq.

certain scenarios, the CM cost function gives rise to global optima that are inadequate in terms of channel inversion.

B. Correntropy-based Criterion

The generalized correlation function proposed in [5] is an attractive concept, since it brings together two interesting features: first, the possibility of taking into account the time structure of signals and, second, the use of statistical information in an extensive manner by encompassing methods derived from ITL, e.g., the kernel estimators, which effectively allows the use of higher-order statistics (HOS). This correlation function or, simply, correntropy, can be defined as $V_y(m) = E\left[K_\sigma(y(n) - y(n - m))\right]$ or estimated through the sample mean:

$$\hat{V}_y(m) = \frac{1}{N - m + 1} \sum_{n=m}^{N} K_\sigma\left(y(n) - y(n - m)\right), \quad (2)$$

where $K_\sigma(\cdot)$ denotes a kernel function, σ is the kernel size, N is the size of the data window used to estimate correntropy and m is the lag being considered. As stated in [5], it is possible to perform blind equalization for temporally-dependent data by imposing a matching between source and equalizer output correntropies, which gives shape to the following criterion:

$$\min_{\mathbf{w}} J_{corr}(\mathbf{w}) = \min_{\mathbf{w}} \sum_{m=1}^{P} \left(V_s(m) - V_y(m)\right)^2, \quad (3)$$

where $V_s(\cdot)$ and $V_y(\cdot)$ are the correntropies of the source $s(n)$ and the equalizer output $y(n)$, respectively, and P is the largest considered lag.

Since the CM criterion, Eq. (1), does not explicitly promote the recovery of the source correlation profile, the correntropy criterion (3) presents advantages in non-i.i.d. scenarios. This duality raised a possibility that will be investigated in this work - to incorporate elements of the data sample dependence profile to the CM criterion. This will be done in two ways, the first of which is to modify the MU-CMA [7] algorithm, originally proposed as a technique for performing multiuser detection.

III. MODIFIED MU-CMA

Although originally proposed to perform space-time equalization in the context of wireless communications [1], the MU-CMA can be a model to create an algorithm capable of dealing with signals composed of dependent samples. Basically, the cost function of the novel method is composed of a canonical "CM-like term" and a cross-correlation term, which penalizes deviations from the source time signature, which must be known *a priori*. Thus, the modified MU-CM criterion, here denoted as MMU-CM, can be defined as:

$$\min_{\mathbf{w}} J_{MMU\text{-}CM}(\mathbf{w}) = \min_{\mathbf{w}} E\left[\left(|y(n)|^2 - R_2\right)^2\right]$$
$$+ \rho \sum_{m=0}^{P} \left(R_y(m) - R_s(m)\right)^2, \quad (4)$$

where $R_y(m)$ and $R_s(m)$ are the autocorrelation functions of the signals $y(n)$ and $s(n)$, respectively, and ρ is the weight

of the correlation term with respect to the CM term. The adaptation of the equalizer coefficients can be done through the use of a stochastic steepest descent approach, in which we will consider an N-sample window, with $N \geq P$, and real-valued signals, resulting in the MMU-CM algorithm (MMU-CMA):

$$\mathbf{w}(n + 1) = \mathbf{w}(n) - \mu\left(\Delta_{CM} + \rho\Delta_R\right), \quad (5)$$

where

$$\Delta_{CM} = \frac{1}{N} \sum_{i=n}^{n+N-1} \left(|y(i)|^2 - R_2\right) \mathbf{x}(i)y(i);$$

$$\Delta_R = \sum_{m=0}^{P} \left(R_y(m) - R_s(m)\right) \frac{1}{N - m}$$
$$\times \sum_{j=n-N+m+1}^{n} \left(\mathbf{x}(j)y(j - m) + \mathbf{x}(j - m)y(j)\right);$$

$$R_y(m) = \frac{1}{N - m} \sum_{k=n-N+m+1}^{n} y(k)y(k - m).$$

IV. CORRELATION RETRIEVAL CRITERION

A closer look at the criterion defined in (4) reveals that its second term alone is able, in principle, to capture the entire correlation profile of the signal of interest. Hence, we propose a simplified criterion, which will be denoted as Correlation Retrieval (CR):

$$\min_{\mathbf{w}} J_{CR}(\mathbf{w}) = \min_{\mathbf{w}} \sum_{m=0}^{P} \left(R_y(m) - R_s(m)\right)^2. \quad (6)$$

This cost function has a striking similarity with the correntropy-based approach [5], except for the fact that it depends exclusively on second-order statistics to characterize the temporal dependence of samples (though these statistics are squared). Hence, it can serve as an interesting counterpoint to correntropy-based solutions. In a manner similar to that used for the MMU-CMA, the CR algorithm (CRA) can be mathematically described as follows:

$$\mathbf{w}(n + 1) = \mathbf{w}(n) - \mu\Delta_R, \quad (7)$$

where the parameters are the same as those presented in (5).

The proposed MMU-CMA and CRA, in addition to being closely related, establish points of contact with the CM and correntropy-based criteria, as we will see in the next section.

V. POINTS OF CONTACT

The proposed MMU-CMA and CRA can provide a gradual interconnection relating the other two mentioned methods in terms of their cost functions, which will prove useful to clarify their performance potential, as we present bellow.

We start by analyzing the CM (1), MMU-CM (4) and CR (6) criteria. It is clear that their relationship mirrors their origins: the first term of MMU-CM is naturally related to the CM criterion, while the second term to the CR criteria. However, if we compare the CM criterion to the CR (or, alternatively, the second-term of MMU-CM), they seem to be slightly different, but, to a certain degree, both of them make

use of the second-order statistics with respect to the equalizer output $y(n)$. Taking a closer look, if we assume $P=0$ and the instantaneous estimate $R_y(0) \approx y(n)y^*(n) = |y(n)|^2$ to be employed in the CR criterion, the resulting cost function is exactly the same as that of CM, except for the constants R_2 and $R_s(0)$. This brief analysis prompts us to state that these cost functions must be similar, albeit they will not have coincident minima, because $R_s(0)$ is, in general, not equal to R_2. In summary, we can see that the three aforementioned cost functions are related by the following expression:

$$J_{MMU-CM}(\mathbf{w}) = J_{CM}(\mathbf{w}) + \rho J_{CR}(\mathbf{w}) \qquad (8)$$

Considering that the optimal solutions of J_{CM} and J_{CR} will not be coincident, the setting of the parameter ρ will be crucial to the balance of MMU-CM. A possible choice for ρ is $1/(2P+2)$, which equally weighs the J_{CM} and J_{CR} terms, resulting in "intermediate solutions" for MMU-CM.

Interestingly, a useful relationship concerning the correntropy-based cost function can also be derived. If we consider a Gaussian kernel function, the second-order Taylor series expansion around zero can be expressed as:

$$V_y(m) \approx \frac{1}{\sqrt{2\pi}\sigma}\left(1 - \frac{1}{\sigma^2}\left(\sigma_{y(n)}^2 - R_y(m)\right)\right) \qquad (9)$$

where $\sigma_{y(n)}^2$ is the variance of $y(n)$, which was considered here to be equal to $\sigma_{y(n-m)}^2$ for the sake of simplification. The constant multiplicative term $1/(\sqrt{2\pi}\sigma)$ does not affect the minimization problem given by Eq. (3), and, moreover, $\sigma_{y(n)}^2$ is only an energy correction term that constrains the output amplitude. Hence, it is possible to express $V_y(m) \approx -R_y(m)$ and, analogously, $V_s(m) \approx -R_s(m)$. Using these approximations in Eq. (3), the resulting cost function is identical to that of CR approach (or to the second term of Eq. (4)). Again, these similarities indicate a relationship between the cost functions, but their points of minima are not necessarily the same. This will become clearer in next section.

VI. ERROR SURFACE ANALYSIS

In order to study the error surfaces, we assume the equalizer to be a two-tap filter and two distinct scenarios, characterized by signals with and without dependent samples.

A. Dependent Sources

In this scenario, the source is originally composed by $+1/-1$ i.i.d. samples and its dependence is generated by a maximum-phase precoder with transfer function $P(z) = 1 + 1.5z^{-1}$. The channel is a minimum-phase system assumed to be $H(z) = 1 + 0.6z^{-1}$. The chosen parameters were $\rho=10$ for MMU-CM; $P=3$ for MMU-CM, CR and correntropy; and a Gaussian kernel with $\sigma=5$. In all criteria, a batch approach was used with $N=40000$.

In Fig. 1, we present the contours of the CM and the CR cost functions, which reveals some similarities between them, like shape and number of minima - denoted by a plus sign "+" for the CM cost and by a "∗" for the CR cost -, which is in consonance with the analysis exposed in Section V. As expected, their minima do not coincide, since $R_2 \neq R_s(0)$.
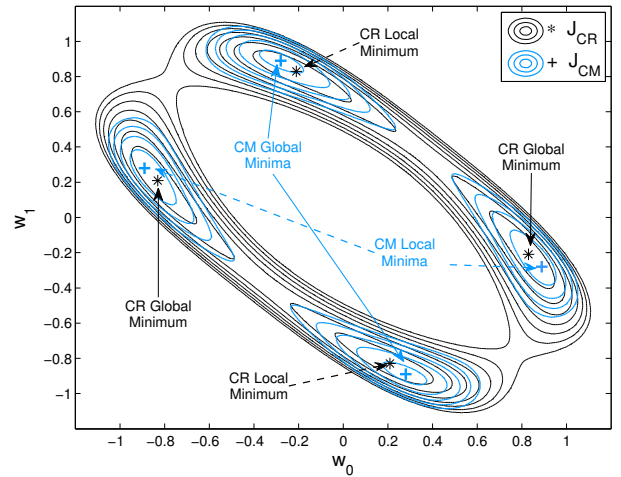

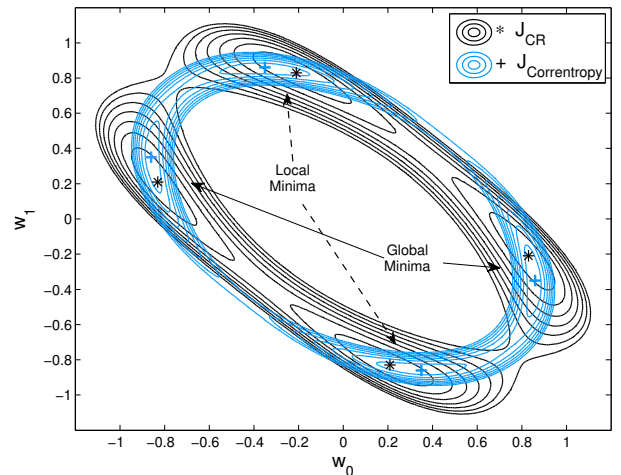Fig. 1.   CM and CR cost function comparison for correlated sources.

Fig. 2.   CR and Correntropy-based cost function comparison for correlated sources.

However, surprisingly, the CM local minima positions are next to CR global minima and vice versa. To understand this, we should consider that other m lags of $R_y(m)$ are used and contribute to the development of the CR global minima, which are able, in this case, to restore the source correlation profile - which is generated by a maximum-phase system, in contrast with the minimum-phase channel. It is also possible to associate all the minima of both CM and CR cost functions to the mean squared error (MSE) solutions for the non-trivial delays of the transmitted signal. While the CM local minima and the CR global minima are associated with the optimum source delay, the CM global minima and the CR local minima are associated with a sub-optimal source delay. For the MMU-CR criteria, it suffices to say that the position of the global minima will depend on ρ. If we consider $\rho > 1/(2P+2)$, the global minima will be located between the correspondent global minima of CR and the local minima of CM.

To investigate the relationship between the CR and the

correntropy-based criteria, we present the contours for their cost functions in Fig. 2. The number of minima is the same and, differently from the CM, their global minima are close, but, on the other hand, their cost functions present a remarkably different shape, which shows us that the approximation $V_y(m) \approx -R_y(m)$ is limited for the situation at hand, being the higher-order even moments non-negligible in view of the nongaussianity of the involved signals.

TABLE I
RESIDUAL ISI CONSIDERING GLOBAL CONVERGENCE.

Source	ISI[dB]			
	CMA	MMU-CMA	CRA	Correntropy
Precoded	-1.5694	-9.2963	-10.0467	-10.0023
i.i.d.	-10.2084	-10.2080	-0.1257	0.2378

These visual analyses cast additional light on the relationship between the focused criteria, but in terms of performance, not much can be straightforwardly inferred. To circumvent this limitation, we show, in Tab. I the residual intersymbol interference (ISI) of the two-tap filter when the algorithms converge to their respective global minima. As expected, the CMA is not able to recover the source correlation profile, performing poorer, while the CRA and the correntropy-based method achieve almost the same level of ISI. The MMU-CMA presents a slight deviation in its global minima with respect to the CRA due to the role of the chosen ρ, implying in a worse performance.

B. Independent Sources

We now analyze the error surfaces for a scenario in which the CM criterion is known to perform well: the case of i.i.d. sources. Considering the same channel of the previous case - without the use of a precoder -, we present the contours of the CM and the CR cost functions in Fig. 3. Similarly to the correlated case, the global minima of the CM are close to the CR local minima; however, they are now in opposite positions, i.e., the minima that most satisfactorily invert the channel are the global minima of the CM. This implies that the CR global minima are not the most adequate solutions in this case. Another interesting point is the greater distance between the CM local minima and the CR global minima. As in the correlated case, both of the minima are associated with MSE solutions for suboptimal delays, but, while the CM implicitly aims at "more independent" output signals, the CR tends to focus on "more dependent" ones, which results in different reference delays and minima positions. This gives rise to multiple local minima in the MMU-CM function, but a more defined global minima for $\rho = 1/(2P+2)$.

Regarding the correntropy-based cost function, it suffices to say that its minima are very close to the correspondent CR minima, in a similar way to Fig. 2. The ISI performances associated with the two-tap global minima of the algorithms are given in Tab. I. As expected, the CM global solution presents the best performance of all.

VII. PERFORMANCE ANALYSIS

The study of two-dimensional filters in the previous section is both simple and clarifying. Real-world tasks may, nonethe-
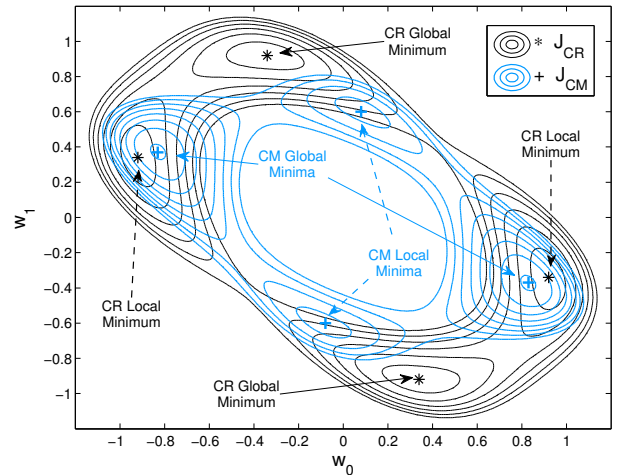

Fig. 3. CM and CR cost function comparison for uncorrelated sources.

less, be associated with more complex channels / equalizers, which raises the question of how general are the conclusions drawn so far in these challenging scenarios. In the following, we will provide the reader with elements of the answer to this question by analyzing two higher-dimensional cases.

In first scenario, we consider the source generated by a duobinary signal submitted to a linear precoder $P(z) = 1 + z^{-1}$ or, in vector notation, $P(z) = [1, 1]$. The signal is distorted by the model of a telephone channel [5] with vector notation $H(z) = [0.04, -0.05, 0.07, -0.21, -0.5, 0.72, 0.36, 0, 0.21, 0.03, 0.07]$. White Gaussian noise is added to the distorted signal, being the SNR equal to 25 dB. We employ a 21-tap equalizer initialized with the center-spike method. The chosen parameters are $\mu_{\text{CMA}}=1e-4$, $\mu_{\text{MMU-CMA}}=6e-5$, $\rho=5$, $\mu_{\text{CRA}}=8e-5$, $\mu_{\text{corr}}=6e-3$; for the MMU-CMA, CRA and correntropy, $P=3$ and $N=10$, so that the variance of their respective equalizer coefficients were similar after convergence. The Gaussian kernel size is calculated in each iteration as the standard deviation of the equalizer output $\sigma = std(y)$. The performances in terms of ISI are illustrated in Fig. 4 for an average of 50 simulations. For this example, it is possible to see that all algorithms seem to converge to suboptimal solutions, given their elevated levels of ISI. However, CRA and the correntropy-based method present better performance and faster convergence, while the CMA is completely unable to capture the correlation profile of the source. The MMU-CMA probably converges to a minimum close to that of the CMA, but there is a deviation caused by the choice of ρ.

For the second case, we consider an alternate mark inversion (AMI) source drawn from the alphabet $[-1, 0, 1]$, an IIR channel $H(z) = 1/(1-0.5z^{-1})$ from [5] and an impulsive additive noise to test the robustness of the algorithms against outliers. The noise is generated according to the Gaussian mixture model presented in [5], with $\epsilon=0.05$, $\sigma_1^2 = 1000\sigma_2^2$ and $\sigma_2 = 0.0114$ (equivalent to a 20 dB SNR). The step-sizes were $\mu_{\text{CMA}}=0.01$, $\mu_{\text{MMU-CMA}}=1e-3$, $\mu_{\text{CRA}}=5e-3$, $\mu_{\text{corr}}=0.08$, and the other parameters were $\rho=15$, $P=4$, $N=10$ and $\sigma=std(y)$,
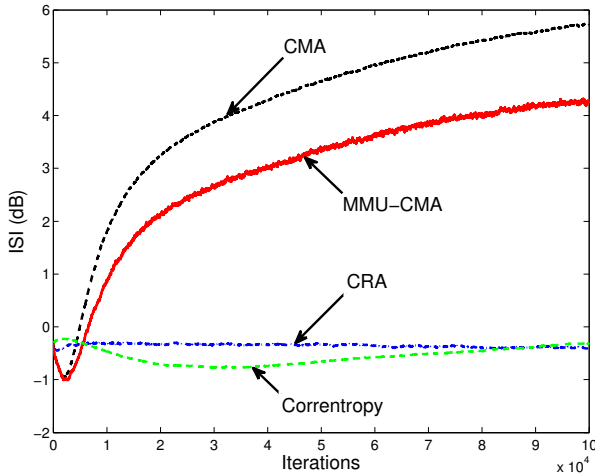
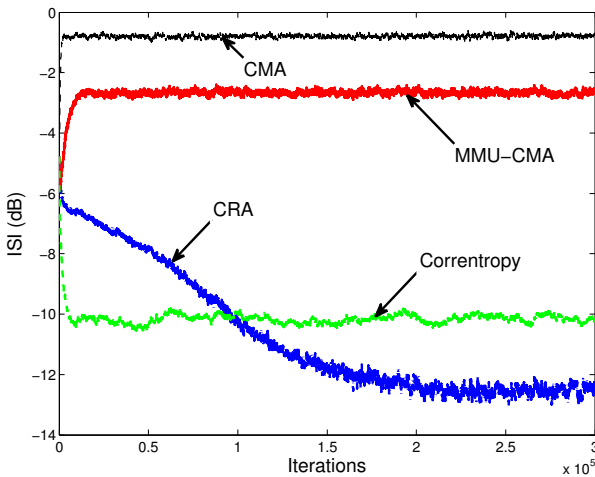Fig. 4.   ISI performance for the telephone channel.

Fig. 5.   ISI performance for the IIR channel.

with the exception of the correntropy-based method, whose data window needed to be larger, with N_{corr}=100, in order to provide more adequate estimates. We considered a 3-tap equalizer initialized with the center-spike method. Fig. 5 displays the ISI curves obtained for this case. Again, the CMA and MMU-CMA presented a worse performance due to suboptimal convergence, being the impulsive noise a likely factor underlying this behavior. On the other hand, CRA and the correntropy-based algorithm achieved a satisfactory performance in inverting the channel. Interestingly, in terms of convergence, the CRA is quite slow, which can be explained, basically, by two main aspects. First, the presence of impulsive noise degrades its performance, but the use of the data window minimizes its effect. Second, the presence of poles in the IIR channel represents additional difficulties for the CRA to correct the source correlation profile only with second-order statistics. In contrast, correntropy is able to deal satisfactorily with these scenarios through the use of HOS and exhibits fast convergence. However, a sufficiently large data window is necessary to provide good estimates, enhancing the required computational complexity. It should be said that, in spite of the

slower convergence, the CRA achieved the lowest ISI level.

VIII. Conclusions

In this work, in order to build a clearer view of the problem of blind deconvolution of dependent sources (signals), we have proposed two new unsupervised algorithms: the MMU-CMA and the CRA. These methods are useful to establish points of contact between the classical CMA and ITL methods like the state-of-the-art correntropy-based algorithm. These points of contact include equivalences and differences present in their cost functions. It has been shown that the CM does not lead to solutions capable of proper channel inversion when the signal of interest is composed of dependent samples, whereas the CRA and the correntropy-based algorithm can perform well, since they consider the time structure of the signals. It is also detailed that the opposite situation is founded when the scenario is characterized by the existence of independent sources. The MMU-CMA is a sort of middle-term solution between the CMA and the CRA, since it is a combination of both of them, but its performance will depend on the setting of the parameter ρ.

The performances of these criteria were analyzed in terms of ISI, which showed that the CRA can serve as an alternative to the correntropy-based methods, since their solutions are relatively close. However, the correntropy-based method presented greater robustness and faster convergence to impulsive noise and larger channels at the cost of the need for a larger data window and, therefore, a higher computational cost. The proposed CRA remains an interesting option due to its simplicity and satisfactory convergence to low ISI levels, even outperforming, in some cases, the correntropy-based algorithm.

In the near future, we intend to seek more elements of comparison between methods for practical scenarios. Another possibility is to explore the connections with BSS methods.

References

[1] J.M.T. Romano, R. Attux, C.C. Cavalcante, R. Suyama, "Unsupervised Signal Processing: Channel Equalization and Source Separation," *CRC Press*, 2010.
[2] D. Godard, "Self-Recovering Equalization and Carrier Tracking in Two-Dimensional Data Communication Systems," *IEEE Trans. on Communications*, v. 28, pp. 1867-1875, 1980.
[3] R. Axford, L. Milstein, J. Zeidler, "The Effects of PN Sequences on the Misconvergence of the Constant Modulus Algorithm (CMA)," *IEEE Trans. on Signal Processing*, v. 46(2), pp. 519-523, 1998.
[4] J. LeBlanc, K. Dogançay, R. Kennedy, R. Johnson, "Effects of Input Data Correlation on the Convergence of Blind Adaptive Equalizers," *IEEE International Conference on Acoustic, Speech and Signal Processing*, v. 3, pp. 313-316, 1994.
[5] I. Santamaría, P. Pokharel, and J. Principe, "Generalized Correlation Function: Definition, Properties and Application to Blind Equalization," *IEEE Trans. on Signal Processing*, v. 54, pp. 2187-2197, 2006.
[6] A. Neves, C. Wada, R. Suyama, R. Attux, J.M.T. Romano, "An Analysis of Unsupervised Signal Processing Methods in the Context of Correlated Sources," *Lect. Notes in Computer Science*, v. 5441, pp.82-89, 2009.
[7] B.C. Papadias, A.J. Paulraj, "A Constant Modulus Algorithm for Multiuser Signal Separation in Presence of Delay Spread Using Antenna Arrays," *IEEE Signal Processing Letters*, v. 4(6), pp. 178-181, 1997.