2004 IEEE 5th Workshop on Signal Processing Advances in Wireless Communications

Impact of Higher-Order Statistics on Adaptive Algorithms for Blind Source Separation

Charles C. Cavalcante and João Marcos T. Romano

Digital Signal Processing for Communications Lab. (DSPCom)
Department of Communications (DECOM)
State University of Campinas (UNICAMP)
R. Albert Einstein, CEP 13083-970, CP 6001, Campinas Brazil
Emails: {charles,romano}@decom.fee.unicamp.br

Abstract — **The paper is devoted to present an analysis of the impact of higher order statistics (HOS) in adaptive blind source separation criteria. Despite the well known fact that they are necessary to provide source separation in a general framework, their impact on the performance of adaptive solutions is a still open research field. The approach of probability density function (pdf) recovering is used. In order to verify the analysis, two constrained adaptive algorithms are investigated. Namely, the multiuser kurtosis algorithm (MUK) and the multiuser constrained fitting probability density function algorithm (MU-CFPA) are used due to the desired characteristics of different HOS involved in their design. Simulation results are carried out to basis our analysis.**

I. Introduction

Blind source separation (BSS) has been gained increasing attention in the signal processing community due to its wide applicability in many fields such as digital communications, biomedical engineering and financial data analysis among others [1].

From the work by Hérault *et al* in 1985 [2] much effort has been done in order to design proper models and suitable statistical criteria that reflect some known structural properties of the sources [3]. A common characteristic of all those criteria is the use of higher order statistics (HOS) since second order statistics (SOS) are not sufficient to solve the separation problem for general sources [4].

The information-theoretic approach has been introduced by Donoho in [5], who has treated the BSS problem by an entropy minimization view point. Other well known method to solve BSS problems is the use of contrast functions introduced by Comon [6], where a contrast function is a cumulant-based function of the separation filter outputs that is maximized if and only if separation is achieved [4, 3].

Those works have provided important results on the issue of necessary and sufficient conditions to provide perfect separation. Despite the development of techniques that rely directly on HOS cumulants, some single user techniques, such as constant modulus (CM) and Shalvi-Weinstein criteria, have been proposed to BSS in a single-stage and multistage context [7, 8].

Papadias proposed in [3, 8] a source separation approach that is based on the Shalvi-Weinstein criterion. The proposal is called multiuser kurtosis (MUK) and consists on the kurtosis maximization, constrained to an orthogonal global response. It has been a great advance on the field of BSS because it has proved global convergence for an arbitrary number of users, what has not been done so far.

The main characteristic of the cumulant-based criteria, as the MUK one, is that they use a number of HOS that has a trade-off between complexity and performance, it means, the criteria has to be the simplest one able to separate the sources. In this work we analyze the impact of using more cumulants on the criteria to be optimized (or in the contrast function) in terms of their corresponding adaptive algorithms. For this sake, we have used an approach of probability density function (pdf) recovering.

We have previously proposed a source separation criterion based on the estimation of the pdf of the ideally recovered signals [9]. The criterion also takes profit from the MUK approach by considering the constraint over the global response in order to provide correct source separation.

Our objective in this work is to evaluate the differences on adaptive solutions when the algorithm considers only one higher order moment, as in MUK case, or *all* higher order moments as our approach in [9]. This aspect can provide significant improvements on the performance of adaptive BSS algorithms and show some guidelines for design blind source separation criteria.

The rest of the paper is organized as follows. Section II shows the analysis involving higher-order cumulants criteria and their impact on adaptive algorithms. On Section III, two algorithms that use different number of cumulants are described to basis the analysis. Section IV shows the computational simulations and discusses the impacts of the use of more higher-order cumulants on the criteria. Finally, our conclusions are stated on Section V.

II. Higher-Order Cumulants and Adaptive Algorithms

We assume that K independent and identically distributed (i.i.d.) and also mutually independent zero mean discrete sequences $a_k(n)$, $k = 1, \ldots, K$, that share the same statistical properties, are transmitted over a MIMO linear memoryless channel that introduces interuser interference.

If we consider M sensors in the receiver we can represent the received signal at time instant n as

$$\mathbf{x}(n) = \mathbf{H}\mathbf{a}(n) + \mathbf{v}(n), \tag{1}$$

where $\mathbf{a}(n) = \begin{bmatrix} a_1(n) & \cdots & a_K(n) \end{bmatrix}^T$ is the vector of sources, $\mathbf{H}$ is the $M \times K$ channel matrix, $\mathbf{v}(n)$ is the $M \times 1$ vector of additive gaussian noise and $\mathbf{x}(n)$ is the $M \times 1$ vector of received signals.

The received signals are then precessed by the MIMO equalizer given by the matrix $\mathbf{W}(n) = \begin{bmatrix} \mathbf{w}_1(n) & \cdots & \mathbf{w}_K(n) \end{bmatrix}$, which produces a $K \times 1$ vector $\mathbf{y}(n)$ that consists of the estimative of the sources. The receiver output can be mathematically written as

$$\begin{aligned}\mathbf{y}(n) &= \mathbf{W}^H(n)\mathbf{x}(n) = \mathbf{W}^H(n)\mathbf{H}\mathbf{a}(n) + \mathbf{v}'(n) \\ &= \mathbf{G}(n)\mathbf{a}(n) + \mathbf{v}'(n),\end{aligned} \tag{2}$$

where

$$\mathbf{G} = \mathbf{W}^H(n)\mathbf{H} = \begin{bmatrix} \mathbf{g}_1 & \cdots & \mathbf{g}_K \end{bmatrix} = \begin{bmatrix} g_{11} & \cdots & g_{K1} \\ \vdots & \ddots & \vdots \\ g_{1K} & \cdots & g_{KK} \end{bmatrix}_{K \times K}$$

is the global response matrix and $\mathbf{v}'(n) = \mathbf{W}^H(n)\mathbf{v}(n)$ is the filtered noise at the receiver output. Figure 1 depicts the above-described system.

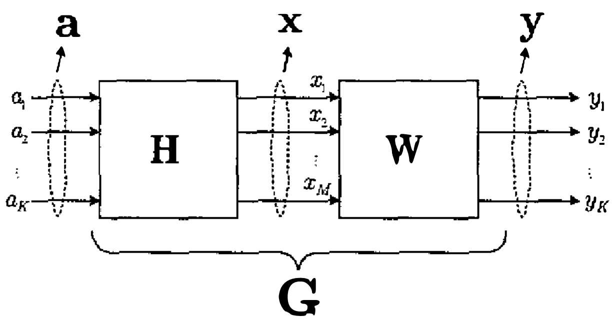

Figure 1: General blind source separation scheme.

But, how to find the matrix $\mathbf{W}$? The answer to this question has been widely investigated through a large number of research teams and published in a great amount of papers.

In [10], one of the pioneers works on blind deconvolution, the approach of pdf matching was inserted. The paper shows that if the pdf of the signal in the input and output of a linear system (channel and equalizer) are equal then the signal of the input is recovered.

After the work [6], much research has been done in order to find cumulant-based contrast functions that when which optimized source separation is achieved. And also, the number of cumulants involved in the contrast function should be chosen to optimize the trade-off complexity $\times$ performance. In [11] some investigation on cumulants was done in order to provide reduction of redundancy in contrast functions.

Based on the Benveniste-Goursat-Ruget (BGR) theorem [10], one may consider that blind source separation aims to equalize the pdfs of the input and output signals of linear systems. If that is achieved, the recovered sources differ from the input one only by a permutation and scale factors [1].

But how is the pdf constructed? In [12], for discrete i.i.d. and zero mean sources, it is proved that the kurtosis can provide correct recovering of the signals. The main analysis was based for closed-form solutions and not for adaptive algorithms.

To evaluate the impact on adaptive algorithms we may use the Gram-Charlier expansion of pdf given by [13]:

$$p_Y(y) = \alpha(y)\left(1 + \sum_{i=3}^{\infty} c_i P_i(y)\right), \tag{3}$$

where $\alpha(y)$ is the pdf of a normalized Gaussian random variable, c_i is the i-th coefficient based on the cumulants up to order i of the pdf of y and P_i terms are the Hermite polynomials defined in terms of the i-th derivative of $\alpha(y)$ as

$$\alpha^{(i)}(y) = (-1)^i \alpha(y) P_i(y). \tag{4}$$

The Gram-Charlier expansion is a general approach of the Edgeworth expansion [14], the later one only writes the expansion in terms of the cumulants in a decreasing form.

Typical Hermite polynomials are:

$$\begin{aligned} P_0(y) &= 1 \\ P_1(y) &= y \\ P_2(y) &= y^2 - 1 \\ P_3(y) &= y^3 - 3y \\ P_4(y) &= y^4 - 6y^2 + 3 \\ P_5(y) &= y^5 - 10y^3 + 15y \\ P_6(y) &= y^6 - 15y^4 + 45y^2 - 15, \end{aligned} \tag{5}$$

and a recursion form of these polynomials is

$$P_{i+1}(y) = yP_i(y) - iP_{i-1}(y) \tag{6}$$

In an adaptive algorithm, the pdf is estimated through the cumulants and also, see Equation (3), by the Hermite polynomials. Actually, if only one cumulant is used as contrast function, the pdf estimaton will be based on this higher-order moment. This is interesting in the steady-state, when the correct solution is achieved, but the convergence performance of the adaptive algorithm can be damaged.

As shown by the Edgeworth expansion , the lower-order cumulants are more significant than the higher-order ones if the pdf is correctly estimated [13]. In fact, the higher-order cumulantes are related to the Hermite polynomials with a lower influence on the estimation, since they carry less information, as it can be seen in Equation (5).

However, in adaptive algorithms, the information provided by the higher-order Hermite polynomials can be very useful in the transient period, when the cumulants cannot be accurately estimated, to improve the pdf estimative and increase the convergence rate.

In order to evaluate these differences we present in the next section two algorithms that use a different number of HOS to perform blind source separation.

III. Two Examples: Kurtosis Maximization and Constrained Fitting PDF Criteria

The presented strategies are based on the well known Shalvi-Weinstein (SW) criterion [12] proposed to the single user (equalization) case. From the SW criterion we know that if the received power (after equalization) is assured to be equal to the transmitted one, it is sufficient to equalize one higher-order moment to achieve equalization, except by a phase rotation. Generalization of this theorem to the multiuser case is done by the insertion of the condition that the recovered sources must be different. Then, the following conditions must hold to assure source separation [3]:

- C1. $a_k(n)$ is i.i.d. and zero mean ($k = 1, \ldots, K$);
- C2. $a_k(n)$ and $a_q(n)$ are statistically independent for $k \neq q$ and have the same pdf;

C3. $|\kappa[y_k(n)]| = |\kappa_a| \quad (k = 1, \ldots, K)$;

C4. $E\{|y_k(n)|^2\} = \sigma_a^2 \quad (k = 1, \ldots, K)$;

C5. $E\{y_k(n) y_q^*(n)\} = 0, \quad k \neq q$.

where $a_k(n)$ is the transmitted sequence by the k-th source, $E\{\cdot\}$ stands for expectation, κ_a is the kurtosis and σ_a^2 is the variance of the transmitted sequence, and $\kappa[\cdot]$ is the kurtosis operator.

In order to prove the sufficiency of the conditions above, we may express the variance and kurtosis of each output as

$$E\{|y_k(n)|^2\} = \sigma_a^2 \sum_{k=1}^{K} |g_{kl}|^2 \tag{7}$$

and

$$\kappa[y_k(n)] = \kappa_a \sum_{k=1}^{K} |g_{kl}|^4 . \tag{8}$$

Then, from Equation (8) and Condition C3 we have

$$\sum_{k=1}^{K} |g_{kl}|^4 = 1, \tag{9}$$

and from Equation (7) and Condition C4 we obtain

$$\sum_{k=1}^{K} |g_{kl}|^2 = 1. \tag{10}$$

Therefore, based on the fact that

$$\sum |g_k|^4 \leq \left(\sum |g_k|^2\right)^2 ,$$

Equations (9) and (10) state that $\mathbf{g}_k$ must be in the form

$$\mathbf{g}_k = \begin{bmatrix} 0 & \cdots & 0 & e^{j\phi_k} & 0 & \cdots & 0 \end{bmatrix}^T , \tag{11}$$

where the single nonzero element can be at any position and ϕ_k is an arbitrary phase rotation. Then, by combining Condition C5 with the noiseless case of Equation (2) we obtain:

$$\mathbf{g}_k^H \mathbf{g}_q = 0, \qquad k \neq q. \tag{12}$$

Equations (11) and (12) dictate the important property that the nonzero position of the solution vectors $\mathbf{g}_k$ and $\mathbf{g}_q$ cannot be the same. Hence, the unique solution that satisfies the problem corresponds to the K solution vectors $\mathbf{g}_k$ be different "Dirac"-type vectors, as given in Equation (11).

In the sequel, we present two algorithms that perform BSS, in according to the conditions given above.

A. Multiuser Kurtosis Algorithm (MUK)

A multiuser algorithm based on the maximization of the kurtosis for blind signals recovering is proposed in [8, 3]. MUK is a constrained criterion that maximizes the kurtosis of the signals, subject to the constraint of normalized global response, it means,

$$\begin{cases} \max\limits_{\mathbf{G}} J_{\mathrm{MUK}}(\mathbf{G}) = \sum\limits_{j=1}^{K} |\kappa[y_k]| \\ \text{subject to: } \mathbf{G}^H \mathbf{G} = \mathbf{I} \end{cases} , \tag{13}$$

where $\mathbf{I}$ is the identity matrix.

The criterion divides the separation task into two parts: the equalization step, that maximizes the kurtosis, and the separation one, that performs the decorrelation of the outputs. For each task, we denote the beamformers by $\mathbf{W}^e$ for equalization part, and $\mathbf{W}$ for the later one. The constraint step is performed by means of a Gram-Schimdt orthogonalization of matrix $\mathbf{W}^e$ [3]. Therefore, these two parts can be, respectively, written in their adaptive versions by [3, 8]:

$$\mathbf{W}^e(n+1) = \mathbf{W}(n) + \mu \,\mathrm{sign}(\kappa_a)\, \mathbf{x}^*(n) \mathcal{Y}(n), \tag{14}$$

where $\mathcal{Y}(n) = \begin{bmatrix} |y_1(n)|^2 y_1(n) & \cdots & |y_K(n)|^2 y_K(n) \end{bmatrix}$ and Equation (14) corresponds to the equalization step. For the orthogonalization one, we have, for the j-th user

$$\mathbf{w}_j(n+1) = \frac{\mathbf{w}_j^e(n+1) - \sum\limits_{l=1}^{j-1} \left(\mathbf{w}_l^H(n+1)\mathbf{w}_j^e(n+1)\right) \mathbf{w}_l(n+1)}{\left\| \mathbf{w}_j^e(n+1) - \sum\limits_{l=1}^{j-1} \left(\mathbf{w}_l^H(n+1)\mathbf{w}_j^e(n+1)\right) \mathbf{w}_l(n+1) \right\|} . \tag{15}$$

The MUK algorithm is summarized in Table 1 [3].

Table 1: MUK algorithm.

1. Initialize $\mathbf{W}(0)$
2. for $n > 0$
3. Obtain $\mathbf{W}^e(n+1)$ from Equation (14)
4. Obtain $\mathbf{w}_1(n+1) = \frac{\mathbf{w}_1^e(n+1)}{\|\mathbf{w}_1^e(n+1)\|}$
5. for $j = 2 : K$
6. Compute $\mathbf{w}_j(n+1)$ from Equation (15)
7. Go to 5
8. Go to 2

This algorithm considers only the fourth order moment (unnormalized kurtosis) to provide source separation. Next section presents an algorithm that uses all higher order moments.

B. Multiuser Constrained Fitting Probability Algorithm (MU-CFPA)

The MU-CFPA [9] is a constrained version of the algorithm proposed for single-user equalization in [15] that has a multiuser version reported in [16]. The original one is based on the estimation of the pdf of an ideally equalized signal at each output, by means of a parametric model that fits the system order and pdf features.

Then, we can construct the criterion in order to minimize the "*distance*" between the desired pdf (ideally equalized one) and the parametric model. Thus, the well known *Kullback-Leibler divergence* (KLD) [17] is used to minimize the divergence between both functions, since both are positive definite functions. The criterion may be written, for the k-th user, as [18]

$$J_{\mathrm{FP}}(\mathbf{w}_k) = D_{p_{Y,\mathrm{ideal}}(y_k) \| \Phi(y,\sigma_r^2)} , \tag{16}$$

where $D_{o||\bullet}$ is the KLD between the pdfs, $p_{Y,\text{ideal}}$ is the pdf of the ideally equalized signal and $\Phi(y, \sigma_r^2)$ is the parametric model given by

$$\Phi(y_k) = \frac{1}{\sqrt{2\pi\sigma_r^2}} \sum_{i=1}^{S} \exp\left(-\frac{\left|\mathbf{w}_k^H(n)\mathbf{x}(n) - a_i\right|^2}{2\sigma_r^2}\right) \Pr(a_i), \quad (17)$$

where σ_r^2 is the variance of each Gaussian in the model, S is the number of symbols in the transmitted constellation and a_i is the i-th symbol from the alphabet of transmitted symbols. It worths mentioning that minimize Equation (16) corresponds to maximize the log-likelihood function [17] and also to find the entropy of y if $\Phi\left(y, \sigma_r^2\right) = p_{Y,\text{ideal}}(y)$ [15, 16, 19].

To assure that the recovered sources be different the orthogonalization procedure of the MUK is inserted in order to cope with the problem of lost users and high steady state error when using an explicit decorrelation term as proposed in [7]. Then, the following constrained criterion has been derived:

$$\begin{cases} \min\limits_{\mathbf{W}} J_{\text{FP}}(\mathbf{W}) = \sum\limits_{k=1}^{K} D_{p_{Y,\text{ideal}}(y)||\Phi(y_k, \sigma_r^2)} \\ \text{subject to: } \mathbf{G}^H\mathbf{G} = \mathbf{I} \end{cases} . \quad (18)$$

and the adaptive version of the algorithm consists in replacing the step 3 in Table 1 by the following expression:

$$\mathbf{W}^e(n+1) = \mathbf{W}(n) - \mu\nabla J_{\text{FP}}\left(\mathbf{W}(n)\right), \quad (19)$$

where $\nabla J_{\text{FP}}\left(\mathbf{W}(n)\right)$ is given by

$$\nabla J_{\text{FP}}\left(\mathbf{W}(n)\right) = \frac{\sum\limits_{i=1}^{S} \exp\left(-\frac{|\mathbf{y}(n) - \mathbf{a}_i|^2}{2\sigma_r^2}\right)\left(\mathbf{y}(n) - \mathbf{a}_i\right)}{\sigma_r^2 \cdot \sum\limits_{i=1}^{S} \exp\left(-\frac{|\mathbf{y}(n) - \mathbf{a}_i|^2}{2\sigma_r^2}\right)} \mathbf{x}^*, \quad (20)$$

where $\mathbf{a}_i$ is the $K \times 1$ vector with the a_i symbol from the transmitted alphabet in all positions.

The resulting algorithm is then called Multiuser Constrained FPA (MU-CFPA) and Table 2 summarizes the dynamic of the algorithm.

Table 2: MU-CFPA.

1. Initialize $\mathbf{W}(0)$
2. for $n > 0$
3. Obtain $\mathbf{W}^e(n+1)$ from Equations (19) and (20)
4. Obtain $\mathbf{w}_1(n+1) = \frac{\mathbf{w}_1^e(n+1)}{\|\mathbf{w}_1^e(n+1)\|}$
5. for $j = 2 : K$
6. Compute $\mathbf{w}_j(n+1)$ from Equation (15)
7. Go to 5
8. Go to 2

The minimization of KLD in Equation (18) is achieved if and only if the two pdf are equal, which consists in matching all statistical moments of both pdfs. So, the MU-CFPA clearly respects the necessary conditions to provide source separation and kurtosis maximization is implicitly comprised in the equalization procedure.

IV. Simulation Results

We consider a simple case of a 2×2 unitary channel matrix [3]

$$\mathbf{H} = \begin{bmatrix} 0.701 + j0.172 & 0.629 + j0.286 \\ -0.274 - j0.634 & 0.159 + j0.704 \end{bmatrix}, \quad (21)$$

for the case of two independent QPSK inputs in a 30 dB signal-to-noise ratio (SNR) environment. The parameters of simulations were: $\mu_{\text{MUK}} = 2 \times 10^{-3}$, $\mu_{\text{MU-CFPA}} = 10^{-2}$, $\sigma_r^2 = 0.1$ and $\mathbf{W}(0) = \mathbf{W}^e(0) = \mathbf{I}$ for both algorithms.

In order to evaluate the performance of both algorithms we use the constant modulus error (CME) defined, for the k-th user, as follows:

$$\text{CME}_k(n) = \left(|y_k|^2 - R\right)^2 . \quad (22)$$

The constant R is related to the power of the transmitted constellation. In our case we will assume a normalized power, it means, $R = 1$. It is important to mention that the step sizes were chosen as the highest ones that allow to reach the lowest CME for both algorithms.

Figure 2 shows the evolution of the mean CME (from both users) of the two algorithms. We can observe that the MU-CFPA outperforms the MUK in terms of convergence rate and reaches almost the same final steady state error (about -30 dB). The convergence of the MU-CFPA is reached about of 250 symbols and the MUK converges around the 3000 symbols. This behavior has also been observed in a wide range of simulations for different channels and signal-to-noise ratios.

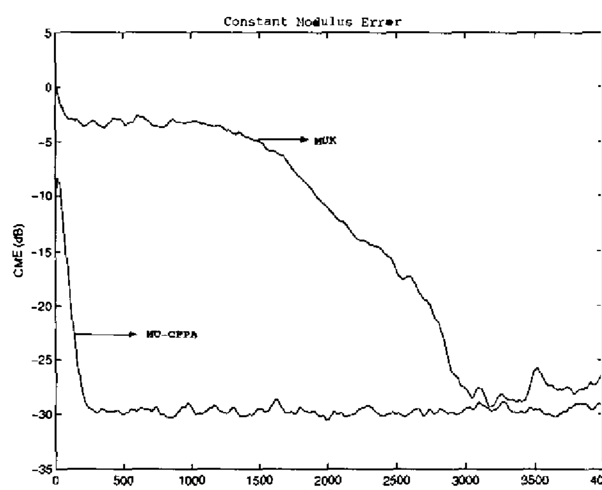

Figure 2: CME evolution for both algorithms.

Further, we can observe that a phase rotation is inserted in the MUK case. This is not observed in the MU-CFPA, due its characteristic of phase recovering [9, 15, 16] that is assured when *all* higher-order moments are used. Figure 3 illustrate the 10%

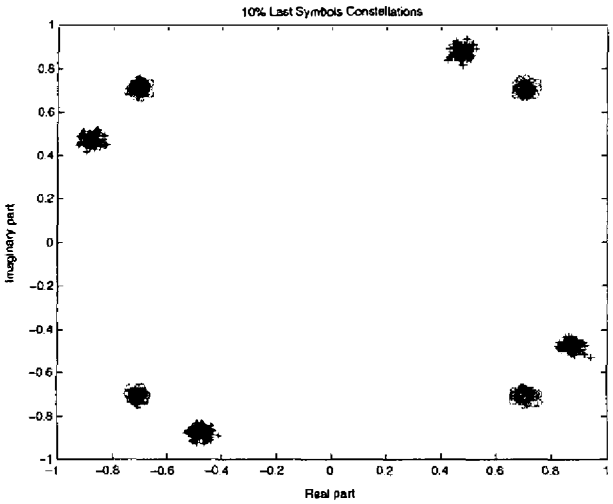

Figure 3: Constellation of the 10% last symbols of one output from both algorithms (MUK – + in black; MU-CFPA – o in gray).

last symbols constellations from both algorithm, so that the phase recovering capability of the MU-CFPA can be observed.

V. Conclusions and Future Works

We have presented some tools for the analysis of the use of higher order moments in adaptive blind source separation algorithms.

The analysis is based on the probability density function matching approach and on the expansion of the pdf in the Gram-Charlier one, that uses the Hermite polynomials which carry information about data during the adaptation procedure.

In order to evaluate the used analytical tools we have studied two adaptive algorithms of the literature, namely the multiuser kurtosis algorithm (MUK) and the multiuser constrained fitting pdf algorithm (MU-CFPA). Those algorithms respect the assumed hypothesis and they use contrast functions with a different number of higher-order moments.

The main feature we have observed is the improvement on the speed of convergence with a small increase in the computational complexity, in the particular case of the MU-CFPA.

A natural extension to this work is to find an optimum point for the trade-off complexity × performance and extend the comparisons with other adaptive algorithms that optimize contrast functions with more than one higher-order cumulants.

References

[1] A. Hyvärinen and E. Oja, "Independent Component Analysis: Algorithms and Applications," *Neural Networks*, vol. 13, no. 4-5, pp. 411–430, 2000.

[2] J. Hérault, C. Jutten, and B. Ans, "Détection de grandeurs primitives dans un message composite par une architecture de calcul neuromimétique en apprentissage non supervisé," in *Actes du Xéme Colloque GRETSI*, Nice, France, Mai 1985, pp. 1017–1022.

[3] C. B. Papadias, "Globally Convergente Blind Source Separation Based on a Multiuser Kurtosis Maximization Criterion," *IEEE Transactions on Signal Processing*, vol. 48, no. 12, pp. 3508–3519, December 2000.

[4] S. Haykin, Ed., *Unsupervised Adaptive Filtering*, ser. (Series on Adaptive and Learning Systems for Signal Processing, Communications and Control). John Wiley & Sons, 2000, vol. I: Source Separation.

[5] D. Donoho, *On Minimum Entropy Deconvolution*. Academic Press, 1981, pp. 565–608.

[6] P. Comon, "Independent Component Analysis: A New Concept?" *Signal Processing*, vol. 36, no. 3, pp. 287–314, April 1994.

[7] C. B. Papadias and A. J. Paulraj, "A Constant Modulus Algorithm for Multiuser Signal Separation in Presence of Delay Spread Using Antenna Array," *IEEE Signal Processing Letters*, vol. 4, no. 6, pp. 178–181, June 1997.

[8] C. B. Papadias, *Blind Separation of Independent Sources Based on Multiuser Kurtosis Optimization Criteria*. John-Wiley & Sons, 2000, vol. 2, ch. 4, pp. 147–179.

[9] C. C. Cavalcante, F. R. P. Cavalcanti, J. C. M. Mota, and J. M. T. Romano, "A Constrained Version of Fitting PDF Algorithm for Blind Source Separation," in *Proceeding of IEEE Signal Processing Advances for Wireless Communications (SPAWC 2003)*, Rome, Italy, June, 15-18 2003.

[10] A. Benveniste, M. Goursat, and G. Ruget, "Robust Identification of a Nonminimum Phase System: Blind Adjustment of a Linear Equalizer in Data Communications," *IEEE Transactions on Automatic Control*, vol. Vol. AC-25, no. No. 3, pp. 385–399, June 1980.

[11] J.-P. Nadal and N. Parga, "Redundancy Reduction and Independent Component Analysis: Conditions on Cumulants and Adaptive Approaches," *Neural Computation*, vol. 9, pp. 1421–1456, 1997.

[12] O. Shalvi and E. Weinstein, "New Criteria for Blind Deconvolution of Nonminimum Phase Systems (Channels)," *IEEE Transactions on Information Theory*, vol. Vol. 36, no. No. 2, pp. 312–321, March 1990.

[13] S. Haykin, *Neural Networks: A Comprehensive Foundation*, 2nd ed. Prentice Hall, 1998.

[14] J.-L. Lacoume, P.-O. Amblard, and P. Comon, *Statistiques d'Ordre Supérieur pour le Traitement du Signal*, ser. (Traitement du Signal). Paris: Masson, 1997.

[15] C. C. Cavalcante, "Neural Prediction and Probability Density Function Estimation Applied to Blind Equalization," Master's thesis, Federal University of Ceará (UFC), Fortaleza-CE, Brasil, February 2001.

[16] C. C. Cavalcante, F. R. P. Cavalcanti, and J. C. M. Mota, "Adaptive Blind Multiuser Separation Criterion Based on Log-Likelihood Maximisation," *IEE Electronics Letters*, vol. 38, no. 20, pp. 1231–1233, September 2002.

[17] S. Haykin, Ed., *Unsupervised Adaptive Filtering*, ser. (Series on Adaptive and Learning Systems for Signal Processing, Communications and Control). John Wiley & Sons, 2000, vol. II: Blind Deconvolution.

[18] C. C. Cavalcante, F. R. P. Cavalcanti, and J. C. M. Mota, "A PDF Estimation-Based Blind Criterion for Adaptive Equalization," in *Proceedings of IEEE Int. Symposium on Telecommunications (ITS 2002)*, Natal, Brazil, September 2002.

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