EUSIPCO 2013 1569745029

USING GENERIC ORDER MOMENTS FOR SEPARATION OF DEPENDENT SOURCES WITH LINEAR CONDITIONAL EXPECTATIONS

Cesar F. Caiafa[1] *and Ercan E. Kuruoğlu*[2]

[1]Instituto Argentino de Radioastronomía (CCT La Plata, CONICET),
C.C.5, 1894 Villa Elisa, Buenos Aires, ARGENTINA.
phone: + (54) 221 482 4903, fax: + (54) 221 425 4909, email: ccaiafa@iar.unlp.edu.ar

[2]Istituto di Scienza e Tecnologie dell'Informazione - CNR,
Via Moruzzi, 1, I-56124 Pisa, ITALY.
phone: + (39) 050 315 3128, fax: + (39) 050 315 2810, email: ercan.kuruoglu@isti.cnr.it

ABSTRACT

In this work, we approach the blind separation of dependent sources based only on a set of their linear mixtures. We prove that, when the sources have a pairwise dependence characterized by the *linear conditional expectation (LCE) law*, i.e. $E[S_i|S_j] = \rho_{ij}S_j$ for $i \neq j$, with $\rho_{ij} = E[S_iS_j]$ (*correlation coefficient*), we are able to separate them by maximizing or minimizing a Generic Order Moment (GOM) of their mixture defined by $\mu_p = E[|\alpha_1 S_1 + \alpha_2 S_2|^p]$. This general measure includes the higher order as well as the fractional moment cases. Our results, not only confirm some of the existing results for the independent sources case but also they allow us to explore new objective functions for Dependent Component Analysis. A set of examples illustrating the consequences of our theory is presented. Also, a comparison of our GOM based algorithm, the classical FASTICA and a very recently proposed algorithm for dependent sources, the Bounded Component Analysis (BCA) algorithm, is shown.

1. INTRODUCTION

We approach the problem of separating n sources from a set of m linear instantaneous mixtures without using any information about the mixing coefficients. Written in a matrix form we have the following model:

$$\mathbf{x}(\tau) = \mathbf{A}\mathbf{s}(\tau), \tag{1}$$

where $\mathbf{x}(\tau) \in \mathbb{R}^m$ is a column vector containing the *mixtures* (measurements), $\mathbf{s}(\tau) \in \mathbb{R}^n$ is a column vector containing the source signals (*sources*) and $\mathbf{A} \in \mathbb{R}^{m\times n}$ is the *mixing matrix* containing the mixing coefficients. The parameter τ is an index that can be related to the position in time or space (pixel index) depending on the application. Then the objective is to estimate the source vector $\mathbf{s}(\tau)$ using only the observations $\mathbf{x}(\tau)$.

Blind Source Separation (BSS) algorithms have been developed during the last two decades based on different kinds of assumptions on the sources (see [9] for an up to date review of BSS algorithms). When sources are statistically independent, the problem can be solved in the sense that sources can be identified up to scale and permutation indeterminacies. This result allowed the development of a sort of Independent Component Analysis (ICA) algorithms which were successfully and widely used in engineering problems.

Unlike in the ICA case, the separation of dependent sources or Dependent Component Analysis (DCA) has not been fully studied in the past and showed more difficulties [1, 4, 6, 5, 3, 16, 19]. Our previously published experimental results revealed that some types of dependent sources, such as the ones found in hyperspectral imaging applications, can be successfully separated by maximizing different measures of non-Gaussianity [6, 5]. However, the theoretical explanation was recently elucidated in [7].

1.1 Notation

We use capital letters to denote scalar random variables and lower case letters for their realizations, for example, S_1, S_2,..., S_n are the random variables associated to the sources which have a *joint probability density funcion (pdf)* denoted by $f_{S_1S_2...S_n}(s_1,s_2,...,s_n)$. Obviously, when sources are independent, the joint pdf factorizes, i.e.

$$f_{S_1S_2...S_n}(s_1,s_2,...,s_n) = f_{S_1}(s_1)f_{S_2}(s_2)...f_{S_n}(s_n), \tag{2}$$

where $f_{S_i}(s_i)$ is called the *marginal pdf* of variable S_i. In this work we are particularly interested in the case of having dependent sources where such a factorization of the joint pdf does not exist. We also define the *conditional pdf* of a random variable S_1 given that $S_2 = x$ as follows:

$$f_{S_1|S_2}(s_1|x) = f_{S_1,S_2}(s_1,x)/f_{S_2}(x). \tag{3}$$

Accordingly, the *first* and *second order conditional expectations* are defined as follows:

$$E[S_1|x] = \int s_1 f_{S_1|S_2}(s_1|x)ds_1, \tag{4}$$

$$E[S_1^2|x] = \int s_1^2 f_{S_1|S_2}(s_1|x)ds_1. \tag{5}$$

Since equations (4) and (5) are functions of x, we use the following notation to express the derivative with respect to x, for example:

$$E'[S_1|x] = \frac{d}{dx}E[S_1|x]. \tag{6}$$

Additionally, in order to avoid conflicts we assume that first and second order derivative of marginal pdf $f'_{S_i}(x)$ and $f''_{S_i}(x)$ always exist for every $x \in (-\infty,+\infty)$.

2. GENERAL STRATEGY FOR DCA

When the matrix $\mathbf{A}$ is full-column rank (equal number or more sensors than sources), we are able to exactly estimate each of the sources by using linear combinations of the mixtures upon permutation and scaling indeterminacies [9]. Therefore, a reasonable strategy used in the ICA context is to search in the space of coefficients, for the points for which each one of the sources is separated. In other words, we need to analyze the behavior of the *mixture random variable X* defined as:

$$X = \alpha_1 S_1 + \alpha_2 S_2 + ... + \alpha_n S_n. \tag{7}$$

Then, we say that variable S_i is separated from the mixture when all coefficients are zero except α_i, i.e. $\alpha_i = 1$ and $\alpha_j = 0$ for every $j \neq i$.

In order to discriminate between a single source compared to any linear combination of two or more sources (keeping the variance fixed), many objective functions have been proposed in the context of ICA. For example, it is known that we can search for the local minima of the Shannon entropy keeping the variance constant because of a classic result from Information Theory: the entropy of a sum of independent variables is larger than the entropy of individual variables [15, 11, 13, 18]. Other objective functions have been proposed for ICA as the case of higher order cumulants for which 4th order cumulant or *kurtosis* is a particular case [8, 15, 14, 20], the convex perimeter for bounded sources [10], L^2-distance non-Gaussianity measure [4], least absolute end-point (LAE) [17], and others. Recently, based on the BCA criterion of [10], a Bounded Component Analysis (BCA) algorithm was proposed in [12] by Erdogan which showed to successfully separate a class of highly correlated dependent sources.

3. GENERIC ORDER MOMENTS

The family of objective functions commonly used for ICA includes higher order moments [8] and, more recently, fractional lower order moments [21, 2]. All these measures can be treated by defining the Generic Order Moment (GOM) of order p:

$$\mu_p \equiv E[|X|^p] = \int_{-\infty}^{+\infty} |x|^p f_X(x) dx. \tag{8}$$

Let us consider now the case of having only two dependent sources i.e. $n = 2$ which are normalized ($E[S_i] = 0$ and $E[S_i^2] = 1$). Since we need to preserve the variance of the mixture, here the constraint is: $\alpha_1^2 + \alpha_2^2 + 2\rho\alpha_1\alpha_2 = 1$, with $\rho = E[S_1 S_2]$ being the correlation coefficient between sources. Then we can parametrize the coefficients by using only one parameter t:

$$\alpha_1(t) = t, \tag{9}$$

$$\alpha_2(t) = -t\rho + \sqrt{t^2(\rho^2 - 1) + 1}. \tag{10}$$

Thus, the GOM associated to the mixture variable $X(t) = \alpha_1(t)S_1 + \alpha_2(t)S_2$ depends explicitly on the parameter t, i.e.

$$\mu_p(t) = E[|X(t)|] = E[|\alpha_1(t)S_1 + \alpha_2(t)S_2|^p]. \tag{11}$$

If the GOM is locally maximum or minimum at $t = 0$, it means that we can separate source S_2 by detecting the point where the derivative of GOM of any linear combination sources is zero (stationary points). The following theorem proves that the GOM is a valid objective function because it has stationary points at $t = 0$.

Theorem 1. ***Stationarity of the GOM measure:*** *Given two zero-mean and unit-norm source variables S_1 and S_2, the GOM of order p, $\mu_p(t) = E[|X(t)|^p]$, has a stationary point at $t = 0$ if the Linear Conditional Expectation (LCE) law, defined by $E[S_1|S_2] = \rho S_2$, is held where $\rho = E[S_1 S_2]$ is the correlation coefficient between sources.*

Proof. We need to prove that the derivative of $\mu_p(t)$ is zero at $t = 0$. By using the chain rule of the derivative, we obtain:

$$\mu_p'(t) = \frac{\partial \mu_p}{\partial \alpha_1}\alpha_1'(t) + \frac{\partial \mu_p}{\partial \alpha_2}\alpha_2'(t). \tag{12}$$

Now, using the fact that $\frac{\partial \mu_p}{\partial \alpha_i} = \int |x|^p \frac{\partial f_X(x)}{\partial \alpha_i} dx$ (i=1,2), taking into account that $\alpha_1'(0) = 1$, $\alpha_2'(0) = -\rho$ and by inserting the partial derivatives $\frac{f_X(x)}{\partial \alpha_i}$ $(i = 1, 2)$ as obtained in Appendix A.1 into eq. (12), we finally obtain:

$$\mu_p'(0) = \int_{-\infty}^{+\infty} |x|^p \underbrace{\left[-\left(f_{S_2(x)} E[S_1|x]\right)' + \rho\left(x f_{S_2}(x)\right)'\right]}_{=0 \text{ if } E[S_1|x]=\rho x \text{ (LCE hypothesis)}} dx, \tag{13}$$

which demonstrates that GOM has a stationary point at $t = 0$. □

In order to determine if the separation of sources is attained at a maximum or a minimum of the objective function, we need to evaluate the second order derivative as follows. Using the fact that the second order derivative of the pdf of the mixture variable is (see Appendix A.2):

$$\left.\frac{d^2 f_X(x)}{dt^2}\right|_{(t=0)} = \left(f_{S_2}(x)E[S_1^2|x]\right)'' + (1 - 3\rho^2) f_{S_2}(x) + x(1 - 5\rho^2) f_{S_2}'(x) - \rho^2 x^2 f_{S_2}''(x), \tag{14}$$

and taking into account the following results:

$$\int \left(f(x)E[S_1^2|x]\right)'' |x|^p dx = p(p-1)\int f(x)E[S_1^2|x]|x|^{p-2} dx,$$

$$\int x|x|^p f'(x) dx = -(p+1)\mu_p,$$

$$\int x^2|x|^p f''(x) dx = (p+2)(p+1)\mu_p,$$

we finally arrive at:

$$\mu_p''(0) = p(p-1)\int |x|^{p-2} f_{S_1}(x) E[S_1^2|x] dx - p\mu_p\left(1 + \rho^2(p-2)\right), \tag{15}$$

Thus, by evaluating equation (15) we are able to determine if the separation is attained at a maximum or a minimum of the GOM. We note that it depends on the distribution of sources and on the second order conditional expectation $E[S_1^2|x]$ function. It is noted that having independent sources is a particular case where the LCE law also holds since $E[S_1|S_2] = E[S_1] = 0$ and $E[S_1 S_2] = E[S_1][S_2] = 0$.

4. THREE SIMPLE DATA SET EXAMPLES

In order to illustrate our results we consider here three simple examples of dependent sources whose scatter plots are shown in Fig. 1:

1. *Data set A: Uncorrelated but dependent sources*: We consider here two sources S_1 and S_2 generated as follows:

$$S_1 = N_1 N_2, \tag{16}$$
$$S_2 = N_2, \tag{17}$$

where N_1 and N_2 are independent non-Gaussian random variables with $E[N_1] = E[N_2] = 0$ and $E[N_1^2] = E[N_2^2] = 1$. We see that S_1 and S_2 are highly dependent but are uncorrelated because $\rho = E[S_1 S_2] = E[N_1 N_2^2] = E[N_1]E[N_2^2] = 0$. The first order conditional expectation is zero, i.e. $E[S_1|S_2] = E[N_1]S_2 = 0$. We also compute the second order conditional expectation which is $E[S_1^2|S_2] = E[N_1^2 N_2^2|N_2] = S_2^2 E[N_1^2] = S_2^2$.
Then, equation (15) is reduced to:

$$\mu_p''(0) = p(p-2)\mu_p, \tag{18}$$

and we conclude that we have a minimum at the separation point ($\mu_p''(0) > 0$) for every $p > 2$ and we have a maximum ($\mu_p''(0) < 0$) for every $p < 2$ ($p \neq 0$).

2. *Data set B: Constrained sources (dependent)*: Motivated by the type of signals observed in the Spectral Unmixing Application of BSS (see [6]), here we generate a special type of sources which are dependent, correlated and constrained to have their sum constant. More specifically, we generate our signals S_1 and S_2 as follows: First, we generate Q independent, nonnegative random variables N_q ($q = 1, 2, ..., Q$) then, we define the following random variables: $U_q = N_q / \sum_{k=1}^{Q} N_k$. We note that these signals meet the constraint $\sum_{q=1}^{Q} U_q = 1$ which is exactly what is observed in the Spectral Unmixing application because sources are associated to abundances (percentages of a particular material). Now, we define our sources by normalizing two of these constrained sources, i.e.:

$$S_1 = (U_1 - \mu_{U_1})/\sigma_{U_1}, \tag{19}$$
$$S_2 = (U_2 - \mu_{U_2})/\sigma_{U_2}. \tag{20}$$

It is not hard to prove that these sources meet the LCE law since $E[S_1 S_2] = \rho = -1/(Q-1)$ and $E[S_1|S_2] = \rho S_2 = -S_2/(Q-1)$.

3. *Data set C: Highly correlated sources*: Here, the sources were constructed using rows of pixels of a one-megapixel digital photo of a man (shown in Fig. 1). We construct the source S_1 by concatenating all the rows of this image starting at the first row and, the source S_2 by using the same concatenation but starting at row number thirty. As the scatter plot shows in Figure 1, these signals are highly dependent (and correlated) because of the high correlation of neighbor pixels in real-world images.

5. SOURCE SEPARATION EVALUATION AND COMPARISON WITH OTHER ALGORITHMS

We consider here the simplest case of having two source signals and two mixtures which correspond to the linear model

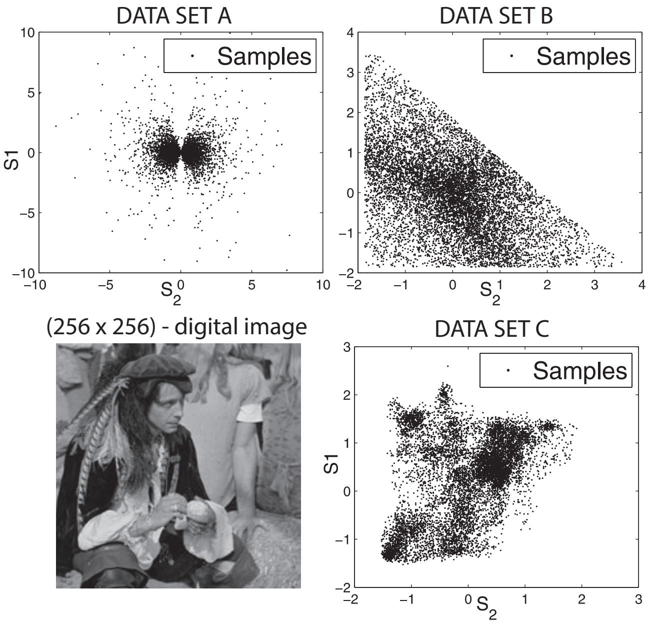

Figure 1: Scatter plots (S_1 versus S_2) of dependent sources data sets.

of equation (1) with a square mixing matrix $\mathbf{A} \in \mathbb{R}^{2\times 2}$. We developed a simple algorithm for source separation by maximizing (or minimizing) the GOM μ_p. As usual, in order to simplify the search of the maximum (or minimum), we first apply a whitening filter, i.e. $\mathbf{y}(\tau) = \mathbf{T}\mathbf{x}(\tau)$ where $\mathbf{y}(\tau)$ is the whitened data ($E[\mathbf{y}\mathbf{y}^T] = \mathbf{I}$) and the filter matrix is given by $\mathbf{T} = \Lambda^{-\frac{1}{2}}\mathbf{U}^T$ with Λ and $\mathbf{U}$ being the diagonal matrix of singular values and the matrix of singular vectors of the covariance matrix $\mathbf{C_{xx}} = E[\mathbf{x}\mathbf{x}^T]$ respectively.

After data is whitened, we can estimate sources by using the following parameterization which maintain the variance fixed:

$$\hat{s}(\tau, \theta) = \cos(\theta)\mathbf{y_1}(\tau) + \sin(\theta)\mathbf{y_2}(\tau). \tag{21}$$

We can quickly estimate the GOM by using the average over a set of T_s samples, i.e,

$$\mu_p(\theta) \approx \sum_{\tau=1}^{T_s} |\hat{s}(\tau, \theta)|^p, \tag{22}$$

which gives a linear complexity ($\mathscr{O}(T_s)$) in terms of the number of samples. The search of a maximum (minimum) can be efficiently done by using a Newton type iteration as follows:

$$\theta_{k+1} = \theta_k \pm \frac{\mu_p'(\theta_k)}{|\mu_p''(\theta_k)|}, \tag{23}$$

where $\mu_p'(\theta_k)$ and $\mu_p''(\theta_k)$ are the first and second order derivative respectively, which can be also computed very fast with linear complexity ($\mathscr{O}(T_s)$).

5.1 Experiments

We applied our algorithm based on GOM with $p = 1.5, 2.5, 10$ to randomly mixed signals generated according to the models of Data set A, Data set B (with parameter $Q = 4$) and Data set C. In all the cases we considered signals length

$T_s = 65536$. In Fig. 2, typical plots of generic order moments with $\beta = 1.5, 2.5, 10$ computed on whitened data for the two sources case, with a random matrix A for Data set B and Data set C in the whole range of the parameter θ. It is noted that, for Data set B we need to search for maxima when $p = 1.5$ or $p = 10$ and minima when $p = 2.5$. On the other hand, for Data set C, we need to search for maxima when $p = 1.5$ and minima when $p = 2.5$ or $p = 10$. It is also noted that for Data set C the local extrema are very smooth which indicate that their associated positions are harder to determine compared to Data set B. It is important to note that only real sources, and their inverted versions (shifted by π), produce local maxima (minima).

In Table 1, the results of our experiments are shown and compared to the results obtained by applying the FASTICA algorithm [14] and the BCA algorithm developed by Erdogan [12]. We used the implementation of FASTICA available at http://research.ics.tkk.fi/ica/fastica/ and applied it by using the two most popular nonlinearities: cubic (u^3) and hyperbolic tangent ($\tanh(u)$). To measure the performance we use the standard Signal to Interference Ratio (SIR) which is defined as $\text{SIR}_i = -10\log_{10}(Var(\hat{s}_i - s_i))$ and the results were averaged over fifty Monte Carlo simulations. It is important to highlight that, for Data set A and B, the best performances were obtained with our algorithm using $p = 1.5$ (SIR=49.67dB) and $p = 10$ (SIR=59.18dB), respectively. For Data set C, the best result was obtained by using the BCA algorithm (SIR=56.62dB) while the results of using our algorithm are also acceptable (SIRs are higher than 23dB). It is interesting to note that the performance of FASTICA (u^3) on Data set A was not bad (SIR=29.94) and this is because in this data set the sources are uncorrelated. On the other hand, when sources are correlated the performance of FASTICA deteriorates significantly and the BCA algorithm was not capable to separate sources in the Data set A and B because they do not meet with the support of sources hypothesis assumed in [12].

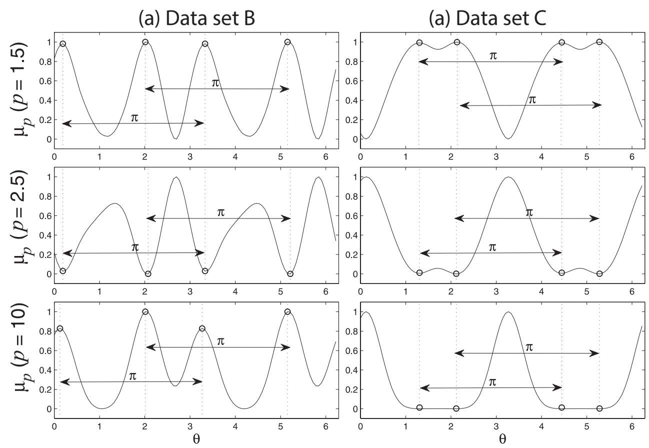

Figure 2: GOM μ_p versus the parameter θ for $p = 1.5$ (top), $p = 2.5$ (middle) and $p = 10$ (bottom) computed for Dataset B (left) and Dataset C (right).

Table 1: Separation performance by using GOM (with $p = 1.5, 2.5, 10$), FASTICA [14] and BCA algorithm [12].

	Mean SIR (dB)		
	Dataset A	Dataset B	Dataset C
FASTICA (u^3)	29.94	15.77	7.34
FASTICA ($\tanh(u)$)	4.74	11.91	7.34
BCA	6.76	18.35	**56.62**
GOM ($p = 1.5$)	**49.67**	44.72	37.12
GOM ($p = 2.5$)	41.23	28.37	33.47
GOM ($p = 10$)	17.14	**59.18**	23.25

6. CONCLUSIONS

We have introduced GOM as a valid objective function for ICA and DCA. These measures are very attractive since they are easily computed based on ergodic averaging with linear complexity ($\mathcal{O}(T_s)$) which makes them useful compared to more sophisticated information theoretic measures such as Renyi or Shannon Entropy which usually requires higher complexity cost. For example, Renyi or Shannon measures computed by using kernel methods have quadratic complexity ($\mathcal{O}(T_s^2)$) [13]. Besides, our theoretical framework opens the possibility to approach new separation problems where sources are allowed to be dependent. Although we have theoretically analyzed the two sources case only, our criteria can be extended also to the general case of having more than two sources but the mathematical treatment would become intricate. Another potential extension of our theory is to generalize it the case of complex signals.

We have provided enough simulation results to validate our theoretical results. Additionally we have compared the performance of our DCA algorithm based on GOM against the classical FASTICA algorithm and and a very recently proposed algorithm , the Bounded Component Analysis (BCA) algorithm. It is noted that, in BCA, the separation is granted when the convex hull of the sources domain can be written as the cartesian product of the convex hulls of the individual source supports [10] which is a very restrictive assumption. It was experimentally verified that the quality of reconstruction of FASTICA algorithm is deteriorated when sources have some degree of dependence.

Appendix A.1

Taking into account that the pdf of the mixture variable X is (for $\alpha_2 > 0$):

$$f_X(x;\alpha) = \frac{1}{\alpha_2}\int f_{S_1S_2}\left(s_1, \frac{x-\alpha_1 s_1}{\alpha_2}\right) ds_1, \qquad (24)$$

and taking the partial derivatives of the pdf evaluated at $(\alpha_1, \alpha_2) = (0, 1)$ we obtain:

$$\frac{\partial f_X(x)}{\partial \alpha_1} = -\left(f_{S_2}(x)E[S_1|x]\right)',$$
$$\frac{\partial^2 f_X(x)}{\partial \alpha_1^2} = \left(f_{S_2}(x)E[S_1^2|x]\right)'',$$
$$\frac{\partial f_X(x)}{\partial \alpha_2} = -\left(x f_{S_2}(x)\right)',$$
$$\frac{\partial^2 f_X(x)}{\partial \alpha_2^2} = 2f_{S_2}(x) + 4x f'_{S_2}(x) + x^2 f_{S_2}(x),$$
$$\frac{\partial^2 f_X(x)}{\partial \alpha_1 \alpha_2} = 2\left(f_{S_2}(x)E[S_1|x]\right)' + x\left(f_{S_2}(x)E[S_1|x]\right)''.$$

Appendix A.2

Using the chain rule of derivatives we have that

$$\frac{d^2 f_X(x;t)}{dt^2} = \frac{\partial^2 f}{\partial \alpha_1^2}\left(\alpha_1'(t)\right)^2 + 2\frac{\partial^2 f}{\partial \alpha_1 \partial \alpha_2}\alpha_1'(t)\alpha_2'(t) + \frac{\partial^2 f}{\partial \alpha_2^2}\left(\alpha_2'(t)\right)^2 + \frac{\partial f}{\partial \alpha_1}\alpha_1'' + \frac{\partial f}{\partial \alpha_2}\alpha_2''. \quad (25)$$

Using the results of Appendix A.1 and the fact that

$$\alpha_1'(0) = 1,\ \alpha_1''(0) = 0,\ \alpha_2'(0) = -\rho,\ \alpha_2''(0) = \rho^2 - 1; \quad (26)$$

we obtain the desired result of equation (14).

Acknoledgments

We thank Dr. Alper Erdogan for providing his Matlab code with the implementation of the Bounded Component Analysis (BCA) used in his recent paper [12]. This work was developed under the scope of the CONICET project PIP 2012-2014, number 11420110100021.

REFERENCES

[1] L. Bedini, D. Herranz, E. Salerno, C. Baccigalupi, E. Kuruoğlu, and A. Tonazzini. Separation of correlated astrophysical sources using multiple-lag data covariance matrices. *EURASIP Journal on Applied Signal Processing*, 2005:2400–2412, 2005.

[2] S. Bulek and N. Erdol. Blind speech separation using fractional order moments. *Statistical Signal Processing, 2009. SSP'09. IEEE/SP 15th Workshop on*, pages 509–512, 2009.

[3] C. Caiafa, E. Kuruoglu, and A. Proto. A minimax entropy method for blind separation of dependent components in astrophysical images. *Bayesian Inference and Maximum Entropy Methods in Science and Engineering (AIP Conf. Proc. Vol. 872)*, 872:81–88, 2006.

[4] C. Caiafa and A. Proto. Separation of statistically dependent sources using an L2-distance non-Gaussianity measure. *Signal Processing*, 86(11):3404–3420, 2006.

[5] C. Caiafa, E. Salerno, and A. Proto. Blind source separation applied to spectral unmixing: comparing different measures of nongaussianity. *Knowledge-Based Intelligent Information and Engineering Systems*, pages 1–8, 2010.

[6] C. Caiafa, E. Salerno, A. Proto, and L. Fiumi. Blind spectral unmixing by local maximization of non-Gaussianity. *Signal Processing*, 88(1):50–68, 2008.

[7] C. F. Caiafa. On the conditions for valid objective functions in blind separation of independent and dependent sources. *EURASIP Journal on Advances in Signal Processing*, 2012(1):1–17, 2012.

[8] J. Cardoso. High-order contrasts for independent component analysis. *Neural Computation*, 11(1):157–192, 1999.

[9] P. Comon and C. Jutten. *Handbook of Blind Source Separation: Independent Component Analysis and Applications*. 2010.

[10] S. Cruces. Bounded Component Analysis of Linear Mixtures: A Criterion of Minimum Convex Perimeter. *Signal Processing, IEEE Transactions on*, 58(4):2141–2154, 2010.

[11] S. Cruces, A. Cichocki, and S. Amari. The minimum entropy and cumulants based contrast functions for blind source extraction. *Bio-Inspired Applications of Connectionism*, pages 786–793, 2001.

[12] A. Erdogan. A family of bounded component analysis algorithms. *Acoustics, Speech and Signal Processing, 2012. ICASSP 2012. IEEE International Conference on*, pages 1–4, Mar. 2012.

[13] D. Erdogmus, I. Hild, E. Kenneth, and J. Principe. Blind source separation using Renyi's [alpha]-marginal entropies. *Neurocomputing*, 49(1-4):25–38, 2002.

[14] A. Hyvarinen. Fast and robust fixed-point algorithms for independent component analysis. *Neural Networks, IEEE Transactions on*, 10(3):626–634, 2002.

[15] A. Hyvärinen, J. Karhunen, and E. Oja. Independent Component Analysis. 2001.

[16] E. Kuruoglu. Dependent Component Analysis for Cosmology: A Case Study. *Latent Variable Analysis and Signal Separation*, pages 538–545, 2010.

[17] J. Lee, F. Vrins, and M. Verleysen. Blind source separation based on endpoint estimation with application to the MLSP 2006 data competition. *Neurocomputing*, 72(1-3):47–56, 2008.

[18] D. Pham, F. Vrins, and M. Verleysen. On the Risk of Using RÉnyi's Entropy for Blind Source Separation. *Signal Processing, IEEE Transactions on*, 56(10):4611–4620, 2008.

[19] E. Savia, A. Klami, and S. Kaski. Fast dependent components for fMRI analysis. *Acoustics, Speech and Signal Processing, 2009. ICASSP 2009. IEEE International Conference on*, pages 1737–1740, 2009.

[20] V. Zarzoso, R. Phlypo, and P. Comon. A contrast for independent component analysis with priors on the source kurtosis signs. *Signal Processing Letters, IEEE*, 15:501–504, 2008.

[21] D. Zha, T. Qiu, H. Tang, Y. Sun, S. Li, and L. Shen. A New Blind Source Separation Method Based on Fractional Lower Order Statistics and Neural Network. *Advances in Neural Networks–ISNN 2004*, pages 678–683, 2004.