A Minimax Entropy Method for Blind Separation of Dependent Components in Astrophysical Images

Cesar F. Caiafa*, Ercan E. Kuruoğlu† and Araceli N. Proto**

*Lab. de Sistemas Complejos, Fac. de Ingeniería - UBA,
Av. Paseo Colón 850 (1063), 4to. piso, Ala sur, Buenos Aires, Argentina.

†Istituto di Scienza e Tecnologie dell'Informazione - CNR,
Via Moruzzi 1, I-56124 Pisa, Italy.

**Comisión de Inv. Científicas de la Prov. de Buenos Aires,
Av. Paseo Colón 751 (1063), Buenos Aires, Argentina.

Abstract. We develop a new technique for blind separation of potentially non independent components in astrophysical images. Given a set of linearly mixed images, corresponding to different measurement channels, we estimate the original electromagnetic radiation sources in a blind fashion. Specifically, we investigate the separation of cosmic microwave background (CMB), thermal dust and galactic synchrotron emissions without imposing any assumption on the mixing matrix. In our approach, we use the Gaussian and non-Gaussian features of astrophysical sources and we assume that CMB-dust and CMB-synchrotron are uncorrelated pairs while dust and synchrotron are correlated which is in agreement with theory. These assumptions allow us to develop an algorithm which associates the Minimum Entropy solutions with the non-Gaussian sources (thermal dust and galactic synchrotron emissions) and the Maximum Entropy solution as the only Gaussian source which is the CMB. This new method is more appropriate than ICA algorithms because independence between sources is not imposed which is a more realistic situation. We investigate two specific measures associated with entropy: Gaussianity Measure (GM) and Shannon Entropy (SE) and we compare them. Finally, we present a complete set of examples of separation using these two measures validating our approach and showing that it performs better than FastICA algorithm. The experimental results presented here were performed on an image database that simulates the measurements expected from the instruments that will operate onboard ESA's *Planck* Surveyor Satellite to measure the CMB anisotropies all over the celestial sphere.

Keywords: Blind Source Separation (BSS), Dependent Component Analysis (DCA), entropic measures, astrophysical images.
PACS: 01.30.Cc; 02.50.-r; 07.05.Pj; 87.57.Gg; 89.70.+c; 95.75.Mn.

INTRODUCTION

The study of Blind Source Separation (BSS) techniques is a very active area of research since they have proved to be useful for solving a broad set of problems in signal processing such as: speech processing [1], subpixel unmixing in hyperspectral images [2, 3, 4, 5], astrophysical signal processing [6, 7, 8, 9], and others. The purpose of a BSS algorithm is to provide a good estimate of unknown source signals by using, as the only available information, their linearly mixed versions.

Early approaches to BSS were done under the strong assumption of independence of the sources reaching to a wide portfolio of algorithms such as FastICA, JADE, and others (see the book [14] for a complete review), which are usually known as Independent Component Analysis (ICA) algorithms.

Dependent Component Analysis (DCA) has recently arisen as a natural extension of ICA [15, 16], i. e., by considering the case where sources are allowed to be dependent (and possibly correlated). Few algorithms for DCA were investigated in the literature, for example, in [15], the separation is obtained by using the temporal correlation of signals. In [2] and [16] it was shown that minimum entropy criteria remains an useful tool even for dependent sources (which are potentially strongly correlated) while the minimum MI criterion, used by most of ICA algorithms, fails in the separation of dependent sources.

In this paper, we provide a novel approach to the blind separation of astrophysical images which exhibit significant levels of correlation among some of the sources. In particular, we work with synthetic simulations of astrophysical images constructed from radiation maps that will be measured by a satellite in the *Planck Surveyor Satellite* mission[1]. Images will be acquired in several channels and each measure can be considered as a linear mixture of several sources of radiations such as the Cosmic Microwave Background (CMB), the synchrotron (SYN) and the galactic dust (DUST) (see [7] for details). In [6, 7], this problem was approached by using ICA methods, but the results are adversely affected by the non realistic assumption of source independence [8].

DEPENDENT COMPONENT ANALYSIS

Given a set of P input signals s_0, s_1,.., s_{P-1} (potentially non independent) with zero-mean ($E(s_i) = 0$) and unit-variance ($E\left(s_i^2\right) = 1$), a set of M linear mixtures (outputs) x_0, x_1,.., x_{M-1} are determined by $x_i(t) = \sum_{j=0}^{P-1} a_{ij} s_j(t) + n_i(t)$, where the coefficients a_{ij} define the linear mixing process and n_0, n_1,.., n_{M-1} are the sensor noise variables. A matrix representation of this linear process is given, as usual, by:

$$\mathbf{x}(t) = A\mathbf{s}(t) + \mathbf{n}(t) \tag{1}$$

where $\mathbf{s}(t) = [s_0\ s_1\ ...\ s_{P-1}]^T$, $\mathbf{x}(t) = [x_0\ x_1\ ...\ x_{M-1}]^T$ and $\mathbf{n}(t) = [n_0\ n_1\ ...\ n_{M-1}]^T$are $P \times 1$, $M \times 1$ and $M \times 1$ time dependent column vectors respectively, and A is a $M \times P$ fixed matrix with linear independent rows and is called the mixing matrix. We will restrict our present work to the overdetermined case, i.e. , with $M \geq P$.

Notice that, in the case of having low level noise and knowing the mixing matrix A, the source estimates can be obtained through a linear combination of the mixtures, i. e.,

$$\widehat{\mathbf{s}}(t) = A^{\dagger}\mathbf{x}(t) \tag{2}$$

where $A^{\dagger}$ is the $P \times M$ Moore-Penrose matrix inverse (or pseudo inverse). But, as we do not have any information about matrix A, our DCA approach uses a parameterization of this pseudo inverse $A^{\dagger}$ which produces source estimates having unit-variance. It is

[1] Planck Mission website: www.rssd.esa.int/Planck

well known that, in the case of having independent sources, the separation is obtained by finding the P most important local minima of the source entropies [10, 13]. This is a natural consequence of the Central Limit Theorem and it can be formally demonstrated by using the entropy power inequality to the case of the variance constrained to one, and using, as a contrast function the sum of the source estimates entropies (see [10] for a mathematical proof). Moreover, the entropy power inequality can be generalized for certain types of dependent variables as was discovered by S. Takano [11] and studied by O. Johnson [12] which allows us to use the Minimum Entropy criterion as a valid method for the dependent case also.

Following this criterion, we select the appropriate matrix $A^{\dagger}$ which minimizes the entropy of the non-Gaussian source estimates. In this paper, we concentrate on two different entropic measures which are: the Shannon Entropy (SE) and the Gaussianity Measure (GM) defined for a generic zero-mean and unit-variance random variable y as follows:

$$H_{SE}(y) = -\int p_y(y)\log(p_y(y))dy; \qquad H_{GM}(y) = -\int \left[p_y(y) - \Phi(y)\right]^2 dy$$

where $\Phi(y) = \frac{1}{\sqrt{2\pi}}\exp(-\frac{1}{2}y^2)$ is the Gaussian pdf and $p_y(y)$ is the pdf corresponding to the random variable y. Notice that our measure of the Gaussianity is related to the L^2-distance of probability density functions (pdfs). For estimating these entropic measures from a set of N samples: $y(0)$, $y(1)$,.., $y(N-1)$, we use a non-parametric pdf estimation technique, namely the Parzen windows [17, 18].

THE MINIMAX ENTROPY METHOD FOR ASTROPHYSICAL IMAGES

In this section, we describe our DCA approach for the particular case of astrophysical signals where some a priori knowledge about the data can be taken into account.

ASSUMPTIONS: In the *Planck Surveyor Satellite mission*, the onboard sensors will be collecting images from the space at different center frequencies [7]. As was made in [8], we consider a simplified model of equation (1) with $M = 4$ mixtures (x_0, x_1, x_2 and x_3) corresponding to measures at the frequencies: 100GHz, 70GHz, 44GHz and 30GHz; and $P = 3$ sources (s_0, s_1 and s_2) corresponding to Cosmic Microwave Background (CMB), galactic Synchrotron (SYN) and galactic dust (DUST) images. The following specific assumptions are considered:

A1: CMB images are Gaussian while SYN and DUST images are non-Gaussian ones.

A2: In theory, CMB-SYN and CMB-DUST are uncorrelated pairs ($E[s_0 s_1] = E[s_0 s_2] = 0$)

A3: Low noise case is considered, it means, the variance of the additive noise in equation (1) is low. In this case, a good estimate of the mixing matrix A is enough in order to obtain acceptable estimates of sources by using the equation (2). As we show in this work, we qualitatively analyze the robustness to noise of our method.

The assumption A2, which was used before by Bedini et al in [8] for obtaining the separation, is not strictly required by our method but can be used for improving the estimation of CMB as we explain in this paper.

OUR METHOD: As it is usually done in several ICA techniques, a first step in our algorithm is to obtain uncorrelated mixtures. This is known as a Whitening step and a way of do it is by applying the second order statistics method known by Karhunen-Loeve Transformation (KLT), i. e., by obtaining a new set of uncorrelated mixtures $\widetilde{\mathbf{x}} = [\widetilde{x}_0\ \widetilde{x}_1\ \widetilde{x}_2]$ through a linear transformation[2]. Since the whitened mixtures $\widetilde{x}_0$, $\widetilde{x}_1$ and $\widetilde{x}_2$ are uncorrelated ones ($E[\widetilde{x}_0\widetilde{x}_1] = E[\widetilde{x}_0\widetilde{x}_2] = E[\widetilde{x}_1\widetilde{x}_2] = 0$), the way of restricting its linear combination to have unit-variance is to use a set of coefficients lying in the unit norm sphere. In other words, we are allowed to use spheric coordinates which reduces the space of search to have two dimensions (θ_0, θ_1). Accordingly, our source estimates can be obtained through the following equation which depends only on two parameters θ_0 and θ_1: $y(\theta_0, \theta_1) = \widetilde{x}_0 \cos\theta_0 + \widetilde{x}_1 \sin\theta_0 \cos\theta_1 + \widetilde{x}_2 \sin\theta_0 \sin\theta_1$. Therefore, the following two steps define the Minimax entropy method:

Minimum Entropy step: Since non-Gaussian sources are related with local minima of entropy, we seek for the local minima of the entropic measure of $y(\theta_0, \theta_1)$ (SE or GM). The source estimates associated with these local minima correspond to the non-Gaussian sources (SYN and DUST in our case).

Maximum Entropy step: It is well known that the Gaussian density has the maximum entropy for a given variance. Therefore, the source estimate associated with the maximum of the entropic measure of $y(\theta_0, \theta_1)$ (SE or GM), corresponds to the Gaussian source which is, in our case, the CMB.

Then, the source estimates are obtained through a linear transformation on the whitened sources:

$$\widehat{\mathbf{s}} = \widetilde{D}\widetilde{\mathbf{x}} \tag{3}$$

where the rows of the 3×3 matrix $\widetilde{D}$ are obtained one by one applying the previous steps.

Using uncorrelatedness for enhancing CMB estimate

The assumption A2 can be used to estimate the CMB image once SYN and DUST images were obtained and avoiding the Maximum Entropy step[3]. Let us assume that we have already obtained the second (SYN) and third (DUST) rows of the separation matrix $\widetilde{D}$. The matrix $\widetilde{D}$ is related to the covariance matrix of sources as follows: $R_{\widehat{\mathbf{s}}\widehat{\mathbf{s}}} = E[\widehat{\mathbf{s}}\widehat{\mathbf{s}}^T] = \widetilde{D}\widetilde{D}^T$. Since $E[\widehat{s}_0\widehat{s}_1] = E[\widehat{s}_0\widehat{s}_2] = 0$ (by assumption A2), it means that the first row of matrix $\widetilde{D}$ must lie in the subspace which is orthogonal to the second and third rows and we can calculate it directly by using, for example, the Gram-Schmidt procedure.

---

[2] Notice that, under the low noise assumption (A3) the resulting dimensionality of the data after the KLT is 3, i. e., one eigenvalue of the covariance matrix is nearly zero and its associated eigenvector can be discarded.

[3] This is useful since, in our experiments, we have noted that the minima (SYN and DUST) are always spiky and easy to detect while the local maximum (CMB) is more difficult to be accurately detected.

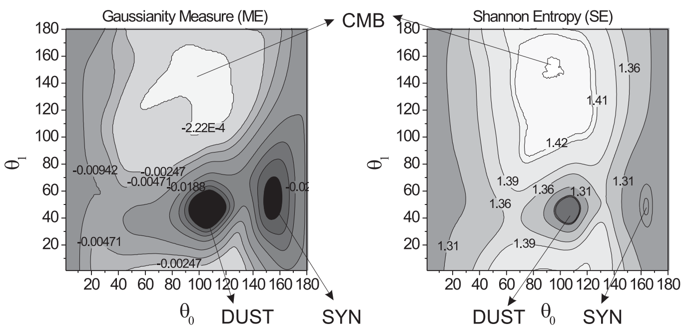

FIGURE 1. 2D-countour plot of GM and SE versus spheric coordinates (θ_0, θ_1)

EXPERIMENTAL RESULTS

Noiseless case

We have synthetically generated the four mixtures by using equation (1) with zero-noise ($\mathbf{n} = \mathbf{0}$) and a selected matrix A according to [6]. We have used images of 256×256 pixels (65536 samples) for CMB, SYN and DUST which were supplied by the Italian Planck team (see [7] for a description on the generation of images). In a pre-processing step, the sources were forced to have zero-mean and unit-variance as it is usual in BSS applications [13]. In the Figure 2, the original sources (top-left) and the synthesized mixtures (bottom) are shown.

The actual matrix $\widetilde{D}$ and its estimates obtained through our algorithm (using the MG and SE measures) are:

$$\widetilde{D} = \begin{bmatrix} -0.36 & 0.62 & 0.70 \\ 1.00 & -0.05 & 0.02 \\ 0.04 & 0.85 & -0.53 \end{bmatrix}$$

$$\widetilde{D}_{MG} = \begin{bmatrix} -0.29 & 0.66 & 0.69 \\ 0.90 & -0.27 & -0.35 \\ -0.07 & 0.71 & -0.70 \end{bmatrix}; \quad \widetilde{D}_{SE} = \begin{bmatrix} -0.28 & 0.67 & 0.69 \\ 1.0 & -0.02 & 0.02 \\ 0.03 & 0.72 & -0.69 \end{bmatrix} \tag{4}$$

In the Figure 1, the corresponding 2D-contour plots for the GM and SE entropic measures are shown for this particular example. Note that the positions of the minima (SYN and DUST) and the maximum (CMB) are clearly identified.

We have obtained the corresponding source estimates ($\widehat{\mathbf{s}} = \widetilde{D}\widetilde{\mathbf{x}}$) using the matrices in (4) and we have calculated the corresponding Signal To Interference Ratio (SIR)[4]

[4] As usually, SIR is defined for each source estimate $\widehat{s}$ in terms of the error variance as follows: SIR$= -10\log_{10}(var(s-\widehat{s}))$

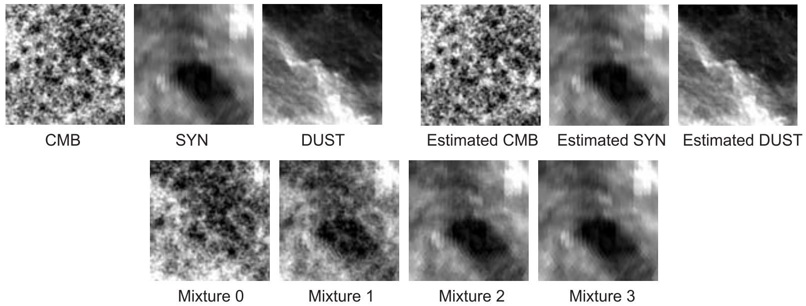

FIGURE 2. Original CMB, SYN and DUST sources (top-left); the four available mixtures (bottom) and the sources estimates using SE (top-right).

as a measure of the separation performance. The obtained SIR levels are shown in the TABLE 1 (Patch number 2). In the Figure 2 (top-right), the source estimates obtained with the SE measure are shown. Note that they are visually identical to the original sources.

Additionally, we have applied the Minimax Entropy algorithm and the FastICA[5] algorithm to a collection of 15 different sets of mixtures corresponding to 15 simulated patches from the whole sky map. In order to find the local minima of the entropic measure (GM and SE), we have used an iterative gradient descend algorithm[6]. The first row of matrix $\widetilde{D}$ was calculated as the unit-norm vector which is orthogonal to the second and third rows as was explained before. In the TABLE 2, the results are shown, note that most of the sources were successfully recovered with our method (a SIR level greater than 8dB can be considered as a successfully recovery). Notice also that SE leads to better separation results than MG and FastICA (higher SIRs).

Robustness against noise

In this section, the sensitivity to noise of the matrix $\widetilde{D}$ estimation technique is analyzed. The surface of the SE was calculated for different levels of Signal To Noise Ratio (SNR) (see Figure 3). It was verified that the positions of the local minima are not dramatically moved away from their original positions as the SNR is decreased (with SNR ≥ 20dB).

---

[5] The ICALAB software package [19] was used for obtaining the FastICA results.

[6] In order to avoid the detection of false local minima we propose to repeat many times the search of a local minimum and, after every local minimum is detected, we proceed with a deflation technique in order to eliminate it avoiding to fall in it again in other cycle search. Finally, we should keep only the two most important minima (the smallest ones).

TABLE 1: Results for DCA with GM and SE entropic measures and FastICA. SIR levels under 8dB are highlighted. The percentage of images sucesfully separated (SIR > 8dB) are: 91%, 93% and 71% for GM, SE and FastICA respectively.

	SIR levels (dB) with GM			SIR levels (dB) with SE			SIR levels (dB) with FastICA		
Patch	**CMB**	**SYN**	**DUST**	**CMB**	**SYN**	**DUST**	**CMB**	**SYN**	**DUST**
1	15.8	10.0	13.9	15.2	13.2	15.6	22.9	13.0	10.2
2	12.0	**7.1**	22.1	13.6	31.9	21.4	8.3	22.1	**6.4**
3	23.3	**2.4**	17.3	23.1	**2.2**	17.4	25.4	13.6	9.8
4	9.8	15.4	8.8	10.3	13.8	9.8	8.7	**6.9**	**2.9**
5	12.5	15.2	20.6	19.7	23.6	20.1	17.5	17.9	22.8
6	14.9	11.5	14.5	18.4	13.0	16.1	20.8	14.8	10.7
7	16.5	11.7	9.6	16.0	14.5	10.3	16.8	15.6	**3.0**
8	42.3	**6.4**	15.6	52.1	**6.7**	15.3	**-1.9**	**5.6**	**7.6**
9	**7.8**	14.7	8.6	11.8	24.1	9.3	**5.0**	**2.6**	**1.5**
10	14.2	13.3	13.8	20.8	22.2	19.4	19.8	16.9	27.0
11	15.8	10.1	14.0	15.3	13.2	14.7	22.7	13.5	10.3
12	20.7	26.6	17.8	21.0	22.5	24.4	28.2	19.1	9.6
13	8.7	11.4	17.4	10.6	**6.2**	12.6	**4.1**	**1.0**	**4.8**
14	11.9	18.6	21.3	13.4	23.6	16.9	13.2	8.5	15.3
15	12.4	15.2	21.0	19.6	23.8	19.9	17.5	17.9	22.8
Mean	**15.9**	**12.6**	**15.7**	**18.7**	**17.0**	**16.2**	**15.3**	**12.6**	**11.0**

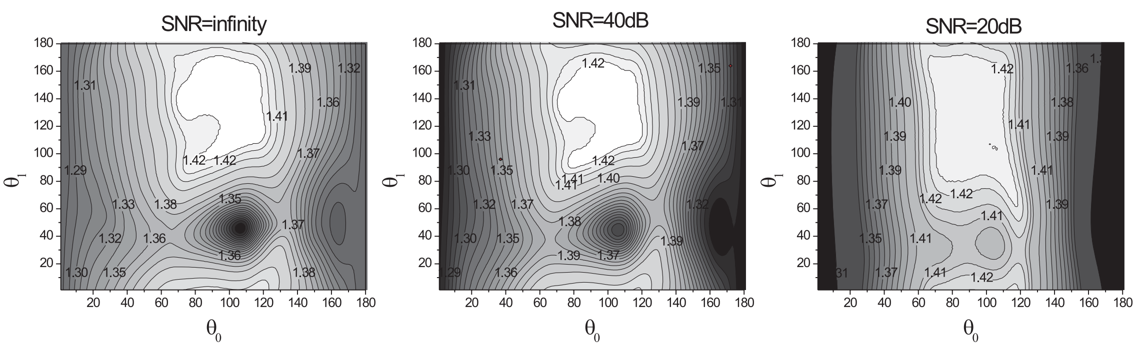

FIGURE 3. SE 2D-contour plot for different SNR levels: ∞, 40dB and 20dB

CONCLUSIONS

A novel method for the separation of astrophysical images from their linear mixtures is introduced which is based on the determination of two minima and one maximum of an entropic measure over the space of the separation parameters (θ_0, θ_1). We have investigted two entropic measures, namely, the Gaussianity Measure (GM) and the Shannon Entropy (SE). While both measures (GM and SE) proved to be useful for obtaining the separation, better results are obtained by using SE. The presented methods are able to separate dependent sources, which is a siginificant development over classical ICA techniques which can separate only independent sources. Our technique was demonstrated to be reasonably robust to low level additive Gaussian noise.

ACKNOWLEDGEMENTS

C. Caiafa thanks to the Jaynes Foundation and the MaxEnt 2006 Conference organizers for their hospitality. This work was supported by Facultad de Ingenieria, Universidad de Buenos Aires, Argentina (Peruilh fellowship and project UBACyT 2003-2007 I050). We are indebted to the Planck teams in Bologna and Trieste, Italy, for supplying us with the maps we used for our simulations.

REFERENCES

1. O. Yilmaz and S. Rickard, "Blind source separation of speech mixtures via time-frequency masking", IEEE Trans. Sig. Proc., **52**, 1830-1847, (2004).
2. C. Caiafa, E. Salerno, A.N. Proto and L. Fiumi, "Dependent Component Analysis as a tool for blind Spectral Unmixing of remote sensed images", EUSIPCO 2006, Florence, Italy, Sept. 4-8, (2006).
3. J. M. P. Nascimento, and J. M. Bioucas Dias. "Does Independent Component Analysis Play a Role in Unmixing Hyperspectral Data?". IEEE TGRS, **43**, 175-187, (2005).
4. N. Keshava and J. Mustard, "Spectral unmixing", IEEE Sig. Proc. Mag., **19**, 44-57,(2002).
5. Naoko Kosaka, Kuniaki Uto, and Yukio Kosugi, "ICA-Aided Mixed-Pixel Analysis of Hyperspectral Data in Agricultural Land", IEEE GRS Lett., **2**, 220-224, (2005).
6. C. Baccigalupi, L. Bedini, C. Burigana, G. De Zotti, A. Farusi, D. Maino, M. Maris, F. Perrotta, E. Salerno, L. Toffolatti, A. Tonazzini, "Neural Networks and Separation of Background and Foregrounds in Astrophysical Sky Maps", Mon. Not. R. Astron. Soc., **318**, 769-780, (2000)
7. E.E. Kuruoglu, L. Bedini, M.T. Paratore, E. Salerno, A. Tonazzini, "Source separation in astrophysical maps using independent factor analysis", Neural Networks, **6**, 479-491, (2003).
8. L. Bedini, D. Herranz, E. Salerno, C. Baccigalupi, E. Kuruoglu, A. Tonazzini, "Separation of correlated astrophysical sources using multiple-lag data covariance matrices", Eurasip J. on Appl. Sig. Proc., **2005**, 2400-2412, (2005).
9. J. Delabrouille, J. F. Cardoso, G. Patanchon, "Multidetector multicomponent spectral matching and applications for cosmic microwave background polarization from satellite astrophysical maps", Mont. Not. R. Astron. Soc., **354**, 55-70, (2004)
10. S. Cruces, A. Cichocki, S-I. Amari, "The Minimum Entropy and Cumulants Based Contrast Functions for Blind Source Extraction", IWANN 2001 Granada, Spain, June 13-15, (2001)
11. S. Takano, “The inequalities of Fisher information and entropy power for dependent variables,” in Proc. 7th Japan-Russia Symp. Prob. Theory and Math. Stat., Tokyo, 26-30 July 1995, S., 460-470, (1996)
12. O. Johnson, "A conditional entropy power inequality for dependent variables", IEEE Trans. Inf Theory, **50**, 1581-1583, (2004).
13. A. Hyvärinen, J. Karhunen and E. Oja, "Independent Component Analysis", J. Wiley & Sons, New York, (2001).
14. A. Cichocki and S. Amari. Adaptive Blind Signal and Image Processing: Learning Algorithms and Applications. J. Wiley & Sons, Chichester UK, (2002).
15. A. Cichocki, T. Rutkowski, A. K. Barros, "Blind Extraction of Temporally Correlated but Statistically Dependent Acoustic Signals", IEEE NNSP ’2000, Sydney, (2000).
16. C. F. Caiafa and A. N. Proto, "Separation of statistically dependent sources using an L^2-distance non-Gaussianity measure", Sig. Proc (EURASIP), **86**, 11, 3404-3420, (2006).
17. B. W. Silverman, Density Estimation for Statistics and Data Analysis, Chapman and Hall, New York 1985.
18. D. Erdogmus, K. E. Hild, J.C. Principe, M. Lazaro, I. Santamaria, "Adaptive blind deconvolution of linear channels using Renyi’s entropy with Parzen window estimation", IEEE Trans. on Sig. Proc, **52**, 1489-1498, (2004).
19. A. Cichocki, S. Amari, K. Siwek, T. Tanaka et al., ICALAB Toolboxes, http://www.bsp.brain.riken.jp/ICALAB.