ICA ALGORITHM BASED ON SUBSPACE APPROACH.

Ali MANSOUR, Member of IEEE and N. Ohnishi, Member of IEEE.

Bio-Mimetic Control Research Center (RIKEN),
2271-130, Anagahora, Shimoshidami, Moriyama-ku,
Nagoya 463 (JAPAN)
mail:mansour@nagoya.riken.go.jp,ohnishi@sheraton.ohnishi.nuie.nagoya-u.ac.jp
http://www.bmc.riken.go.jp/sensor/Mansour

ABSTRACT

Generally, the blind separation algorithms based on the subspace approach are very slow. In addition, they need a considerable computation effort and time due to the estimation and the minimization of huge matrices.

Previously, we proposed an adaptive subspace criterion to solve the blind separation problem [1]. The criterion has been minimized adaptively using a conjugate gradient algorithm [2]. Unfortunately, the convergence of that algorithm needed more than one hour of computational time using an ultra sparc 30 and "C" code program.

In this paper, we improve that criterion by proposing a new subspace adaptive algorithm. The new algorithm deals with stationary signals. The experimental results show that the convergence of the new algorithm is relatively fast due to the estimation by bloc of the different matrices and the minimization of the cost function using a generalized conjugate gradient method.

1. INTRODUCTION

Since 1985, many researchers have been interested by the blind separation of sources problem (or the Independent Component Analysis "ICA" problem) [3, 4, 5, 6, 7, 8, 9, 10]. According to "blind separation" problem, one should estimate, using the output signals of an unknown channel (i.e. the observed signals or the mixing signals), the unknown input signals of that channel (i.e. sources). The sources are assumed to be statistically independent from each other.

Most of the blind separation algorithms deal with a linear channel model: The instantaneous mixtures (i.e. memory-less channel) and the convolutive mixtures (i.e. the channel effect can be considered as a linear filter). The criteria of those algorithms were generally based on high order statistics [11, 12, 13, 14, 15].

Recently, by using only second order statistics, some subspace methods have been explored to separate blindly the sources in the case of convolutive mixtures [16, 17, 18, 19].

In this paper, we propose a new subspace algorithm which improves the performance of our previous criterion [1]. This new algorithm can be decomposed into two steps: First step, by using only second-order statistics, we reduce the convolutive mixture problem to an instantaneous mixture (deconvolution step); then in the second step, we must only separate sources consisting of a simple instantaneous mixture (typically, most of the instantaneous mixture algorithms are based on fourth-order statistics).

2. MODEL & ASSUMPTIONS

Let $Y(n)$ denotes the $q \times 1$ mixing vector obtained from p unknown sources $S(n)$ and let the $q \times p$ polynomial matrix $\mathbf{H}(z) = (h_{ij}(z))$ denotes the channel effect (see fig. 1).

Generally, the authors assume that the sources are statistically independent from each other and that the filters $h_{ij}(z)$ are causal and finite impulse response (FIR) filters. Let us denote by M the highest degree[1] of the filters $h_{ij}(z)$. In this case, $Y(n)$ can be written as:

$$Y(n) = \sum_{i=0}^{M} \mathbf{H}(i) S(n-i), \tag{1}$$

where $S(n-i)$ is the $p \times 1$ source vector at the time $(n-i)$ and $\mathbf{H}(i)$ is the real $q \times p$ matrix corresponding to the filter matrix $\mathbf{H}(z)$ at time i. Let $Y_N(n)$ (resp. $S_{M+N}(n)$) denotes the $q(N+1) \times 1$ (resp. $(M+N+1)p \times 1$) vector given by:

$$Y_N(n) = \begin{pmatrix} Y(n) \\ \vdots \\ Y(n-N) \end{pmatrix}, \tag{2}$$

$$S_{M+N}(n) = \begin{pmatrix} S(n) \\ \vdots \\ S(n-M-N) \end{pmatrix}. \tag{3}$$

By using $N > q$ observations of the mixture vector, we can formulate the model (1) in another form:

$$Y_N(n) = \mathbf{T}_N(\mathbf{H}) S_{M+N}(n), \tag{4}$$

Dr. N. Ohnishi is also a Prof. in Department of Information Engineering, Nagoya University, Furo-cho, Chikusa-ku, Nagoya 464-01, Japan

[1] M is called the degree of the filter matrix $\mathbf{H}(z)$.

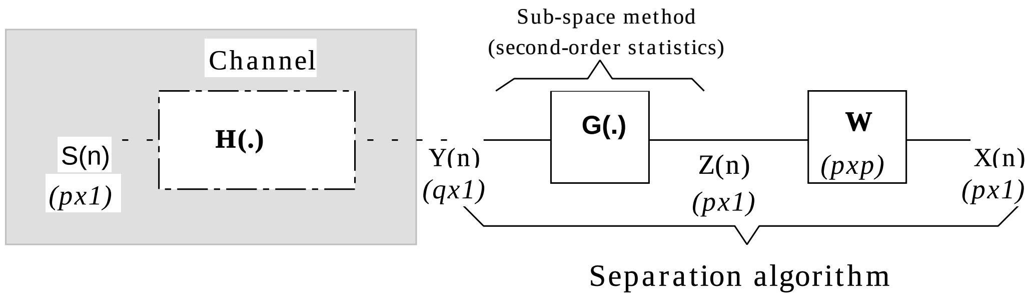


Figure 1: General structure.

where $\mathbf{T}_N(\mathbf{H})$ is the Sylvester matrix corresponding to $\mathbf{H}(z)$. The $q(N+1) \times p(M+N+1)$ matrix $\mathbf{T}_N(\mathbf{H})$ is given by [20] as:

$$\begin{bmatrix} \mathbf{H}(0) & \mathbf{H}(1) & \dots & \mathbf{H}(M) & 0 & \dots & 0 \\ 0 & \mathbf{H}(0) & \dots & \mathbf{H}(M-1) & \mathbf{H}(M) & 0 & \dots \\ \vdots & \vdots & \ddots & \ddots & \ddots & \ddots & \vdots \\ 0 & \dots & 0 & \mathbf{H}(0) & \mathbf{H}(1) & \dots & \mathbf{H}(M) \end{bmatrix}.$$

According to [?], if $H(z)$ is a full rank[2] and a column-reduced matrix (for the definition of column-reduced matrix see [20]), the Sylvester matrix can identify $H(z)$ up to a scalar polynomial filter.

From equation (4), one can conclude that the separation of the sources can be achieved by estimating a $(M+N+1)p \times q(N+1)$ left inverse matrix $\mathbf{G}$ of the Sylvester matrix, which exists if the matrix $\mathbf{T}_N(\mathbf{H})$ has a full rank. In another hand, it was proved in [21] that the rank of $\mathbf{T}_N(\mathbf{H})$ is given by:

$$\text{Rank } \mathbf{T}_N(\mathbf{H}) = p(N+1) + \sum_{i=1}^{p} M_i. \qquad (5)$$

where M_i is the degree of the ith column[3] of $\mathbf{H}(z)$. It is easy to prove using (5) that the Sylvester matrix has a full rank and it is left invertible if each column of the polynomial matrix $\mathbf{H}(z)$ has the same degree and $N > Mp$.

3. CRITERION & ALGORITHM

In a previous paper [1], we present a sub-space algorithm to solve the problem of blind separation of sources for convolutive mixtures. That algorithm was based on the minimization, using the conjugate gradient algorithm, of a subspace criterion which has been based on second-order statistics:

$$\min_{\mathcal{G}} \mathcal{G} \sum_{n=n_0}^{n_1} \mathcal{Y}(n)\mathcal{Y}^T(n)\mathcal{G}^T. \qquad (6)$$

where $\mathbf{G} = (\mathbf{G}_1^T, \dots, \mathbf{G}_{(M+N+1)}^T)^T$ is the estimated left inverse of $\mathbf{T}_N(\mathbf{H})$, $\mathbf{G}_i$ is the ith $p \times q(N+1)$ bloc row of $\mathbf{G}$ and $\mathcal{G} = (\mathbf{G}_1, \mathbf{G}_2, \dots, \mathbf{G}_{(M+N+1)})$ is a $p \times q(N+1)(M+N+1)$ matrix. It has been also shown, in that previous paper [1], that the minimization of the cost function (6) does not give the Moore-Penrose generalized inverse (pseudoinverse) of the Sylvester matrix $\mathbf{T}_N(H)$, but a $(M+N+1)p \times q(N+1)$ matrix $\mathbf{G}$ which satisfies that $\mathbf{G}\mathbf{T}_N(\mathbf{H})$ is a block diagonal matrix:

$$\mathbf{G}\mathbf{T}_N(\mathbf{H}) = \begin{pmatrix} \mathbf{A} & 0 & \dots & \dots & 0 \\ 0 & \mathbf{A} & 0 & \dots & 0 \\ & \ddots & \ddots & \ddots & \\ 0 & & & \ddots & 0 \\ 0 & 0 & \dots & 0 & \mathbf{A} \end{pmatrix}, \qquad (7)$$

where $\mathbf{A}$ is an arbitrary $p \times p$ matrix. It is clear that as the algorithm converges, the estimated sources are instantaneous mixtures (according to a matrix $\mathbf{A}$) of actual sources: in fact using (4) and (7), we find that:

$$\mathbf{G}Y_N(n) = \begin{pmatrix} AS(n) \\ \vdots \\ AS(n-M-N) \end{pmatrix}. \qquad (8)$$

To avoid the spurious solution $\mathbf{G} = \mathbf{0}$ and force the matrix $\mathbf{A}$ to be an invertible matrix[4], it was proposed that the minimization should done subject to the constraint:

$$\mathbf{G}_1\mathbf{R}_Y(n)\mathbf{G}_1^T = \mathbf{I}_p, \qquad (9)$$

where $\mathbf{R}_Y(n) = EY_N(n)Y_N(n)^T$ is the covariance matrix of $Y_N(n)$ and $\mathbf{I}_p$ is a $(p \times p)$ identity matrix. Even if the convergence of that algorithm was attained in small number of iterations (in general case, less than 1000 iterations was needed), but the convergence time is relatively important due to the adaptive minimization of large size matrices.

In this paper, to increase the performance of that criterion in the case of stationary signals, we suggest the following modification of the cost function:

$$\min_{\mathcal{G}} \mathcal{G}\mathbf{A}\mathcal{G}^T, \qquad (10)$$

where $\mathbf{A} = E\mathcal{Y}(n)\mathcal{Y}^T(n)$ is a $q(N+1)(M+N+1) \times q(N+1)(M+N+1)$. One can remark that $\mathbf{A}$

[2] To satisfy those constraints, one must assume that the number of sensors is great than the number of sources $q > p$.

[3] The degree of a column is defined as the highest degree of the filters in this column.

[4] So the separation of the residual instantaneous mixture becomes possible using any algorithm for the separation of instantaneous mixture.

can be evaluated with respect to the covariance matrix[5] $\mathbf{R}_Y = \mathbf{R}_Y(n)$ and it is equal to :

$$\begin{bmatrix} \mathbf{R}_{Y1} & -\mathbf{R}_Y & 0 & 0 & \dots & 0 & 0 \\ -\mathbf{R}_{Y1}^T & 2\mathbf{R}_Y & -\mathbf{R}_{Y1} & 0 & \dots & 0 & 0 \\ 0 & -\mathbf{R}_{Y1}^T & 2\mathbf{R}_Y & -\mathbf{R}_{Y1} & 0 & \dots & 0 \\ \vdots & 0 & \dots & \dots & \dots & 0 & \dots \\ 0 & 0 & \dots & 0 & -\mathbf{R}_{Y1}^T & 2\mathbf{R}_Y & -\mathbf{R}_{Y1} \\ 0 & 0 & \dots & \dots & 0 & -\mathbf{R}_{Y1}^T & \mathbf{R}_Y \end{bmatrix}$$

where $\mathbf{R}_{Y1} = E\mathbf{Y}_N(n)\mathbf{Y}_N^T(n+1)$. Let $\mathbf{B}$ denotes the $q(N+1)(M+N+1) \times q(N+1)(M+N+1)$ matrix:

$$\mathbf{B} = \begin{pmatrix} \mathbf{R}_Y & 0 \\ 0 & 0 \end{pmatrix} \tag{11}$$

Experimentally, $\mathbf{R}_Y$ and $\mathbf{R}_{Y1}$ are estimated, at the beginning of the algorithm, according to the estimation algorithm of [22].

To increase the performances and the convergence speed of the algorithm, the cost function (10) is minimized using a generalized conjugate gradient algorithm, proposed by Chen *et al.* in [23]. That algorithm minimizes the generalized version of Rayleigh's ratio: $f(V) = V^H AV/(V^H BV)$ with respect to a vector (V) (from theoretical point of view, this algorithm can converge in a number of iterations which is less than the dimension of V).

In our case, the cost function (6) must be minimized with respect to a $p \times q(N+1)(M+N+1)$ matrix $\mathcal{G}$. So, let us denote by $\mathcal{G}_i$ the ith row of $\mathcal{G}$, one can verify that the cost function (10) and the constraint (9) can be reevaluated[6] as:

$$\begin{cases} \min_{\mathcal{G}_1} \mathcal{G}_1 \mathbf{A} \mathcal{G}_1^T \\ \text{Subject to} \mathcal{G}_1 \mathbf{B} \mathcal{G}_1^T = 1 \end{cases}$$

and

$$\begin{cases} \min_{\mathcal{G}_1} \mathcal{G}_2 \mathbf{A}_2 \mathcal{G}_2^T \\ \text{Subject to } \mathcal{G}_2 \mathbf{B} \mathcal{G}_2^T = 1 \end{cases}$$

with $\mathbf{A}_2 = \mathbf{A} + \mathbf{B}\mathcal{G}_1^T\mathcal{G}_1\mathbf{B}$. Finally, the source separation of the instantaneous residual mixture is achieved according to the method proposed in [24].

4. EXPERIMENTAL RESULTS

The experimental study shows that for two stationary sources, the convergence of the subspace criterion (10) is attained with less than 300 iterations (see figure 2). The performances are similar to the performances of our previous algorithm [1] but the convergence is obtained in few minutes due to the minimization of the new cost function and the estimation of $\mathbf{A}$ and $\mathbf{B}$ (as described in the previous section).

[5] For stationary signals, the covariance matrix $\mathbf{R}_Y(n)$ is independent of time.

[6] With out less of generality, we will consider just the case of $p = 2$. Anyway, the case $p > 2$ can be easily deduced.

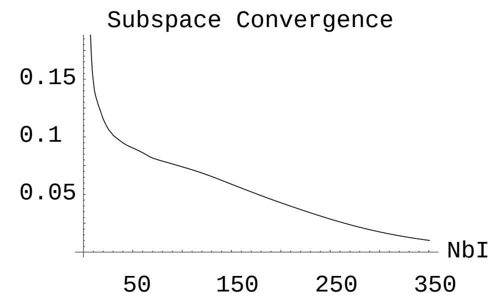


Figure 2: The convergence of the sub-space criterion with respect to the iteration number.

In that experiment, four sensors $q = 4$ and two stationary sources $p = 2$ with an uniform probability density function (pdf) were used. The channel effect $\mathbf{H}(z)$ is considered as a FIR filter of fourth degree $(M = 4)$.

We can see in figure 3 that the objective of first step of the algorithm was achieved, with $G.T_N(H)$ being a block diagonal matrix (where A is a 2×2 matrix, see (7)).

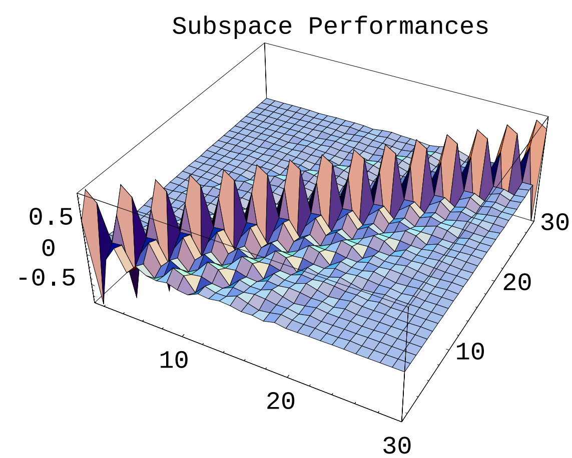
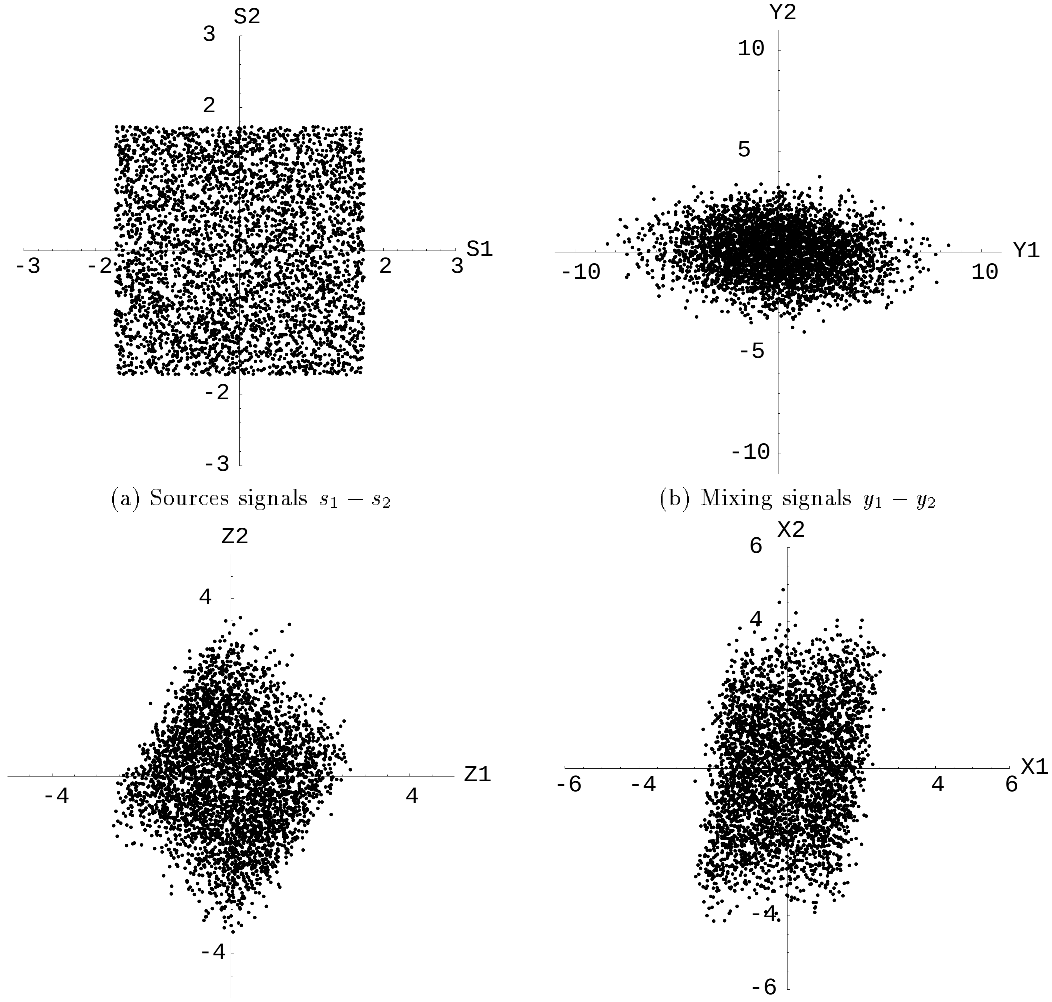


Figure 3: Performance results: $G.T_N(H)$ should be a block diagonal matrix.

Finally, to demonstrate the behavior of our algorithm and its performances, we plot the different signals in their own space, as in figure 4.

In figure 4, we remark that the sources $s_1(n)$ and $s_2(n)$ are statistically independent and so are the estimated signals $x_1(n)$ and $x_2(n)$ (for more information concerning the relationship between the distribution of signals and their statistical relationships with each other, see [25]). In addition, from figure 4 (c) we can say that these signals may be obtained by mixing independent signals with help of an instantaneous mixtures. Finally, we can see the mixing signals, $y_1(n)$ and $y_2(n)$, in the figure 4 (b).



(a) Sources signals $s_1 - s_2$

(b) Mixing signals $y_1 - y_2$

(c) The output of the subspace algorithm: $z_1 - z_2$

(d) Estimated signals $x_1 - x_2$

Figure 4: Experimental results.

5. CONCLUSION

In this paper, we present a blind separation of stationary sources algorithm for convolutive mixtures and based on subspace approach.

This algorithm can be decomposed into two parts: The deconvolution part, using only second order statistics and subspace criterion, and the instantaneous separation of the instantaneous residual mixture using fourth order statistics.

The minimization of the subspace criterion was done using the generalized conjugate gradient algorithm. By consequence, we find that most of the channel parameters can be estimated using only second-order statistics.

The experimental results show that the separation was achieved in few hundred iterations (generally, less than 500 iterations was needed to achieve the subspace deconvolution part). The actual version of the algorithm is relatively fast and we succeeded in separating two stationary sources, with about -20 dB of residual cross-talk. Currently, we are trying to separate non-stationary sources (for example: speech signals).

6. REFERENCES

[1] A. Mansour, A. Kardec Barros, and N. Ohnishi. Subspace adaptive algorithm for blind separation of convolutive mixtures by conjugate gradient method. In *The First International Conference and Exhibition Digital Signal Processing (DSP'98)*, pages I–252–I–260, Moscow, Russia, June 30-July 3 1998.

[2] Z. Fu and E. M. Dowling. Conjugate gradient eigenstructure tracking for adaptive spectral estimation. *IEEE Trans. on Signal Processing*, 43(5):1151–1160, May 1995.

[3] C. Jutten and J. Hérault. Blind separation of sources, Part I: An adaptive algorithm based on a neuromimetic architecture. *Signal Processing*, 24(1):1–10, 1991.

[4] M. Gaeta and J. L. Lacoume. Sources separation without a priori knowledge: the maximum likelihood solution. In L. Torres, E. Masgrau, and M. A. Lagunas, editors, *Signal Processing V, Theories and Applications*, pages 621–624, Barcelona, Espain, 1994. Elsevier.

[5] J. F. Cardoso and P. Comon. Tensor-based independent component analysis. In L. Torres, E. Masgrau, and M. A. Lagunas, editors, *Signal Processing V, Theories and Applications*, pages 673–676, Barcelona, Espain, 1990. Elsevier.

[6] P. Comon. Independent component analysis, a new concept? *Signal Processing*, 36(3):287–314, April 1994.

[7] A. Mansour and C. Jutten. Fourth order criteria for blind separation of sources. *IEEE Trans. on Signal Processing*, 43(8):2022–2025, August 1995.

[8] S. I. Amari, A. Cichoki, and H. H. Yang. A new learning algorithm for blind signal separation. In *Neural Information Processing System 8*, pages 757–763, Eds. D.S. Toureyzky et. al., 1995.

[9] O. Macchi and E. Moreau. Self-adaptive source separation using correlated signals and cross-cumulants. In *Proc. Workshop Athos working group*, Girona, Spain, June 1995.

[10] A. Mansour and C. Jutten. A direct solution for blind separation of sources. *IEEE Trans. on Signal Processing*, 44(3):746–748, March 1996.

[11] C. Jutten, L. Nguyen Thi, E. Dijkstra, E. Vittoz, and Caelen J. Blind separation of sources: An algorithm for separation of convolutive mixtures. In *International Signal Processing Workshop on Higher Order Statistics*, pages 273–276, Chamrousse, France, July 1991.

[12] L. Nguyen Thi and C. Jutten. Blind sources separation for convolutive mixtures. *Signal Processing*, 45(2):209–229, 1995.

[13] A. Mansour and C. Jutten. A simple cost function for instantaneous and convolutive sources separation. In *Actes du XVème colloque GRETSI*, pages 301–304, Juan-Les-Pins, France, 18-21 septembre 1995.

[14] N. Delfosse and P. Loubaton. Adaptive blind separation of convolutive mixtures. In *Proceding of ICASSP*, pages 2940–2943, Atlanta, Georgia, May 1996.

[15] M. Kawamoto, K. Matsuoka, and N. Ohnishi. Blind signal separation of convolved non-stationary signals. In *Proc. International Symposium on Nonlinear Theory and its Applications*, pages 1001–1004, Hawaii, 1998.

[16] D. Gesbert, P. Duhamel, and S. Mayrargue. Subspace-based adaptive algorithms for the blind equalization of multichannel fir filters. In M.J.J. Holt, C.F.N. Cowan, P.M. Grant, and W.A. Sandham, editors, *Signal Processing VII, Theories and Applications*, pages 712–715, Edinburgh, Scotland, September 1994. Elsevier.

[17] A. Gorokhov and P. Loubaton. Second order blind identification of convolutive mixtures with temporally correlated sources: A subspace based approch. In *Signal Processing VIII, Theories and Applications*, pages 2093–2096, Triest, Italy, September 1996. Elsevier.

[18] A. Mansour, C. Jutten, and P. Loubaton. Subspace method for blind separation of sources and for a convolutive mixture model. In *Signal Processing VIII, Theories and Applications*, pages 2081–2084, Triest, Italy, September 1996. Elsevier.

[19] A. Gorokhov and P. Loubaton. Subspace based techniques for second order blind separation of convolutive mixtures with temporally correlated sources. *IEEE Trans. on Circuits and Systems*, 44:813–820, September 1997.

[20] T. Kailath. *Linear systems*. Prentice Hall, 1980.

[21] R. Bitmead, S. Kung, B. D. O. Anderson, and T. Kailath. Greatest common division via generalized Sylvester and Bezout matrices. *IEEE Trans. on Automatic Control*, 23(6):1043–1047, December 1978.

[22] A. Mansour, A. Kardec Barros, and N. Ohnishi. Comparison among three estimators for high order statistics. In *Fifth International Conference on Neural Information Processing (ICONIP'98)*, pages 899–902, Kitakyushu, Japan, 21-23 October 1998.

[23] H. Chen, T. K. Sarkar, S. A. Dianat, and J. D. Brule. Adaptive spectral estimation by the conjugate gradient method. *IEEE Trans. on Acoustics, Speech and Signal Processing*, ASSP-34(2):272–284, April 1986.

[24] A. Mansour, A. Kardec Barros, M. Kawamoto, and N. Ohnishi. A fast algorithm for blind separation of sources based on the cross-cumulant and levenberg-marquardt method. In *Fourth International Conference on Signal Processing (ICSP'98)*, pages 323–326, Beijing, China, 12-16 October 1998.

[25] G. Puntonet, C., A. Mansour, and C. Jutten. Geometrical algorithm for blind separation of sources. In *Actes du XVème colloque GRETSI*, pages 273–276, Juan-Les-Pins, France, 18-21 september 1995.