ELSEVIER

Signal Processing 81 (2001) 2253–2266

SIGNAL PROCESSING

www.elsevier.com/locate/sigpro

A mutually referenced blind multiuser separation of convolutive mixture algorithm

Ali Mansour*

Bio-Mimetic Control Research Center (RIKEN), 2271-130, Anagahora, Shimoshidami, Moriyama-ku, Nagoya 463-0003, Japan

Received 7 June 2000; received in revised form 4 December 2000

Abstract

In this paper, we present a new subspace adaptive algorithm for the blind separation problem of a convolutive mixture. The major advantage of such an algorithm is that almost all the unknown parameters of the inverse channel can be estimated using only second-order statistics. In fact, a subspace approach was used to transform the convolutive mixture into an instantaneous mixture using a criterion of second-order statistics. It is known that the convergence of subspace algorithms is in general, very slow. To improve the convergence speed of our algorithm, a conjugate gradient method was used to minimize the subspace criterion. The experimental results show that the convergence of our algorithm is improved due to the use of the conjugate gradient method.

Keywords: Subspace approach; Second- and higher-order statistics; Sylvester matrix; Blind separation; Convolutive mixture; Conjugate gradient

1. Introduction

Since 1990, the blind separation of sources has been an important issue for the signal processing community. In effect, it can be found in many practical applications and situations (radar control [11], the study of electrocardiogram signals [9], control of a nuclear reactor [12] and the study of seismic signals [38]). This problem was first introduced by Hérault et al. [17], where they proposed a heuristic algorithm based on a biological model [19]. The blind separation problem involves the retrieval of the sources from observations of unknown mixtures of unknown sources [32].

Over the last 15 years, many methods and different algorithms have been proposed to solve this problem in the case of an instantaneous mixture (or memoryless channel) [3,5,6,24,26]. Since 1990, a few methods for source separation have been proposed in the case of convolutive mixtures (i.e. the channel effects can be considered as a linear filter). These methods were generally based on high-order statistics [10,20,25,36]. The major problems of the algorithms based on high-order statistics are the estimation of these statistics and the estimation errors [31].

* Tel.: + 81-52-736-5867; fax: + 81-52-736-5868; http://www.bmc.riken.go.jp/~mansour.
E-mail addresses: mansour@nagoya.riken.go.jp, mansour@ieee.org (A. Mansour).

0165-1684/01/$ - see front matter
PII: S0165-1684(00)00279-6

Nomenclature

$\boldsymbol{A}$	mixing matrix of the residual instantaneous mixture
$\boldsymbol{G}$	left inverse of $\boldsymbol{T}_N(\mathscr{H})$
$\boldsymbol{G}_i$	ith bloc line of $\boldsymbol{G}$
$\mathscr{G}$	another version of $\boldsymbol{G}$
$\boldsymbol{H}(i)$	$q \times p$ real matrix which represents the impulse response of the channel at time i
$\boldsymbol{H}_{cc}$	$q \times p$ non-polynomial matrix
$\mathscr{H}(z) = (h_{ij}(z))$	channel filter ($h_{ij}(z)$ is the filter between the ith source and the jth sensor)
M	degree of the channel
M_i	degree of the ith column of $H(z)$
n	time
N	number of observations
p	number of sources
q	number of sensors
$\boldsymbol{R}_S(m)$	correlation matrix of the sources
$\boldsymbol{R}_Y(m)$	correlation matrix of the observations
$S(s_i)(n)$	vector of the sources (s_i is the ith source)
$S_{M+N}(n)$	giant vector which contains $(M + N + 1)$ vectors of the sources
$\boldsymbol{T}_N(\mathscr{H})$	Sylvester's matrix
$\boldsymbol{W}$	separation matrix of the residual instantaneous mixture
$X(n)$	estimated signals
$Y(n)$	vector of the observations
$Y_N(n)$	giant vector which contains $(N + 1)$ vectors of the observed signals $(Y(n), Y(n-1), \ldots)$
$\mathscr{Y}n$	big matrix formed by the observed signals
$Z(n)$	output of the subspace algorithm

Recently, it has been proven [2,8,13,16,22,27,33,39,40] that the convolutive model can be estimated using only second-order statistics. Most of these methods,[1] in general, are based on subspace theories and approaches. The advantage of subspace methods is that by using only second-order statistics (but more sensors than sources), the sources can be separated (with some assumptions concerning the channel filters) or the convolutive mixture can be identified up to an instantaneous mixture. The subspace methods are highly refined from the theoretical point of view, but in general, the convergence of these algorithms is relatively slow due to the minimization of cost functions containing large matrices.

In Ref. [29], we proposed a subspace algorithm for a convolutive mixture model using the least-mean-square (LMS) algorithm. Unfortunately, that algorithm was very slow due to the minimization, using the LMS algorithm, of a cost function composed of large matrices. In fact, the subspace algorithm requires more than 7000 iterations for convergence and more than several hours of computing time using a sparc ultra 30 and C code. In this paper, we propose another criterion, also based on the subspace approach, which can be

[1] In the decorrelation approaches, the authors consider different assumptions, such as colored signals [8], the system should be strictly dynamic and have some special relation with the minimum phase [22], or the channel should be strictly causal $\boldsymbol{H}(0) = 0$ [39]. On the other hand, the subspace approach generally leads to very elegant algorithms from a theoretical point of view, and is based on a strong theoretical background. It has been developed over many years in control theories [32].

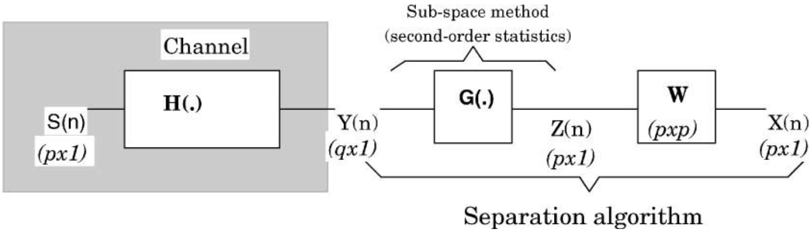

Fig. 1. General structure.

minimized using the conjugate gradient algorithm [7]. In theory, the conjugate gradient algorithm can converge within a few iterations (less than the dimension of the updated vector). The convergence of the proposed method is relatively fast, and may be achieved in less than 1000 iterations and needs less than half an hour of computing time using the same computer.

The algorithm proposed in this paper can be broken down into two steps. First, using only second-order statistics, we reduce the convolutive mixture problem to an instantaneous mixture problem. In the second step, we only separate sources consisting of a simple instantaneous mixture according to the algorithm proposed in Ref. [34] (typically, most of the instantaneous mixture algorithms are based on fourth-order statistics).

2. Channel model

Let us assume that p unknown sources $S(n)$ are statistically independent of each other (this assumption is very common in the blind separation field[2]). In addition, let $Y(n)$ denote the q observed signals (see Fig. 1).

If we consider the mixture to be convolutive, the relationship between the sources and the observed signals can be given by

$$Y(n) = [\mathscr{H}(z)]S(n), \tag{1}$$

where a $q \times p$ polynomial matrix $\mathscr{H}(z) = (h_{ij}(z))$ represents the channel effects, and $h_{ij}(z)$ are assumed to be finite impulse response (FIR) filters.

Let M denote the degree of the filter matrix $\mathscr{H}(z)$, i.e., M is the highest degree of the filters $h_{ij}(z)$ ($\forall 1 \leqslant i \leqslant q$ and $\forall 1 \leqslant j \leqslant p$). $\boldsymbol{H}(i)$ denotes the $q \times p$ real constant matrix corresponding to the impulse response of the channel $\mathscr{H}(z)$ at time i:

$$\mathscr{H}(z) = (h_{ij}(z)) = \sum_{i=0}^{M} \boldsymbol{H}(i) z^{-i}. \tag{2}$$

Eq. (1) can be rewritten as

$$Y(n) = \sum_{i=0}^{M} \boldsymbol{H}(i) S(n-i), \tag{3}$$

[2] This assumption is used in the second step of the proposed algorithm to achieve the separation of the residual instantaneous mixture.

where $S(n-i)$ is the $p \times 1$ source vector at time $(n-i)$. Considering $(N+1)$ observations of the mixture vector $(N > q)$ and using the following notations:

$$Y_N(n) = \begin{pmatrix} Y(n) \\ \vdots \\ Y(n-N) \end{pmatrix} \quad \text{and} \quad S_{M+N}(n) = \begin{pmatrix} S(n) \\ \vdots \\ S(n-M-N) \end{pmatrix}, \tag{4}$$

model (3) can be rewritten as

$$Y_N(n) = \boldsymbol{T}_N(\mathscr{H}) S_{M+N}(n), \tag{5}$$

where the $q(N+1) \times p(M+N+1)$ matrix $\boldsymbol{T}_N(\mathscr{H})$ is the Sylvester matrix corresponding to $\mathscr{H}(z)$. In Ref. [21], the Sylvester matrix is given by

$$\boldsymbol{T}_N(\mathscr{H}) = \begin{bmatrix} \boldsymbol{H}(0) & \boldsymbol{H}(1) & \boldsymbol{H}(2) & \cdots & \boldsymbol{H}(M) & \boldsymbol{0} & \boldsymbol{0} & \cdots & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{H}(0) & \boldsymbol{H}(1) & \cdots & \boldsymbol{H}(M-1) & \boldsymbol{H}(M) & \boldsymbol{0} & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \boldsymbol{0} \\ \boldsymbol{0} & \cdots & \cdots & \cdots & \boldsymbol{0} & \boldsymbol{H}(0) & \boldsymbol{H}(1) & \cdots & \boldsymbol{H}(M) \end{bmatrix}. \tag{6}$$

In the following we will assume the following three assumptions:

H1: The number of sensors is larger than the number of sources, $p < q$. (A method for estimating the number of sources is given in Ref. [2].)

H2: $\mathscr{H}(z)$ is irreducible (Rank$(\mathscr{H}(z)) = p, \forall z$ excluding $z = 0$ but including $z = \infty$).

H3: $\mathscr{H}(z)$ is a column-reduced matrix:
$\mathscr{H}(z)$ can be written as

$$\mathscr{H}(z) = \boldsymbol{H}_{cc} \operatorname{diag}\{z^{-M_1}, \ldots, z^{-M_p}\} + \mathscr{H}_1(z), \tag{7}$$

where M_j denotes the degree[3] of the jth column of $\mathscr{H}(z)$, $\boldsymbol{H}_{cc}$ is a non-polynomial matrix, and $\mathscr{H}_1(z)$ is a polynomial matrix whose degree of the jth column is less than M_j. By definition, $\mathscr{H}(z)$ is reduced by column if and only if $\boldsymbol{H}_{cc}$ is a full-rank matrix.

As long as $p < q$, these assumptions have been shown in Ref. [16] to be realistic (it is easy to verify that if $\mathscr{H}(z)$ is a square column reduced and non-constant matrix, then the rank of $\mathscr{H}(z)$ will be less than p, at least for some z_i such that $\det(\mathscr{H}(z_i)) = 0$). It has been shown in Ref. [4,21] that under the assumptions H2 and H3:

$$^4\text{Rank}(\boldsymbol{T}_N(\mathscr{H})) = p(N+1) + \sum_{i=1}^{p} M_i, \tag{8}$$

as long as $N \geqslant \sum_{i=1}^{p} M_i$. One should note that $p(N+1) + \sum_{i=1}^{p} M_i$ is precisely the number of non-zero columns of $\boldsymbol{T}_N(\mathscr{H})$. In particular, if all the degrees $(M_i)_{i=1,p}$ coincide with M, then, $\boldsymbol{T}_N(\mathscr{H})$ is full column rank if $N \geqslant pM$. Therefore, $\boldsymbol{T}_N(\mathscr{H})$ has a left inverse.

[3] The degree of the jth column of $\mathscr{H}(z)$ equals the maximum degree of the jth column component $h_{ij}(z)$, $\forall 1 \leqslant i \leqslant q$.

[4] In our approach, one should know the number of sources and the filter degree to evaluate the rank of $\boldsymbol{T}_N$ $(\mathscr{H})$ (N and q are known). In the literature, there are many references which emphasize the problems of the estimation of the source number and the degree of the filter, such as[1,2,28,29].

3. Criterion and constraint

Let us assume that the degrees M_i are equal[5] to M:

$$M_i = M \quad \forall i \in \{1, \dots, p\}. \tag{9}$$

Generalizing the method proposed by Gesbert et al. [14] for identification (in the identification problem, the authors assume that they have one source, $p = 1$, and that the source is an independent identically distributed (iid) signal), we propose the estimation of a left inverse matrix of the Sylvester matrix $T_N(H)$ by adaptatively minimizing a cost function.

It is obvious from Eqs. (4) and (5) that the source separation will be achieved by estimating $\boldsymbol{S}_{M+N}(n)$. Consequently, the separation can be performed by estimating a $(M+N+1)p \times q(N+1)$ left inverse matrix $\boldsymbol{G}$ of the Sylvester matrix $\boldsymbol{T}_N(\mathscr{H})$, which exists if the matrix $\boldsymbol{T}_N(\mathscr{H})$ has a full rank.

Assuming that $\boldsymbol{G}$ is the left inverse of $\boldsymbol{T}_N(\mathscr{H})$, we have

$$\begin{aligned} &\boldsymbol{G}\boldsymbol{Y}_N(n) = \boldsymbol{S}_{M+N}(n), \\ &\boldsymbol{G}\boldsymbol{Y}_N(n+1) = \boldsymbol{S}_{M+N}(n+1). \end{aligned} \tag{10}$$

Denoting the ith block row[6] of $\boldsymbol{G}$ by $\boldsymbol{G}_i$ and using Eq. (10), it can easily be proven that

$$\begin{aligned} \mathscr{G}\mathscr{Y}(n) &= (\boldsymbol{G}_1, \boldsymbol{G}_2, \dots, \boldsymbol{G}_{(M+N+1)}) \begin{pmatrix} Y_N(n) & 0 & \dots & 0 \\ -Y_N(n+1) & Y_N(n) & 0 & \vdots \\ 0 & -Y_N(n+1) & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \ddots & -Y_N(n+1) & Y_N(n) \\ 0 & \dots & 0 & -Y_N(n+1) \end{pmatrix}, \\ &= 0. \end{aligned} \tag{11}$$

Here, $\mathscr{G} = (\boldsymbol{G}_1, \boldsymbol{G}_2, \dots, \boldsymbol{G}_{(M+N+1)})$ is a $p \times q(N+1)(M+N+1)$ matrix and $\mathscr{Y}$[7] is a $q(N+1)(M+N+1) \times (N+M)$ matrix defined by the last equation (11). From the same equation, a simple criterion can be derived:

$$\min_{\mathscr{G}} \mathscr{G} \sum_{n=n_0}^{n_1} \mathscr{Y}(n)\mathscr{Y}^{\mathrm{T}}(n)\mathscr{G}^{\mathrm{T}}. \tag{12}$$

The sum operation is added to increase the performance of the experimental results and the robustness of the algorithm (in our experimental study, we used $20 < n_1 - n_0 < 50$).

[5] If this assumption is not satisfied, then, by adopting another parameterization also based on the Sylvester matrix, it is possible to separate the sources [27,28].

[6] $\boldsymbol{G}_i$ is $p \times q(N+1)$ matrix and $\boldsymbol{G} = (\boldsymbol{G}_1^{\mathrm{T}}, \dots, \boldsymbol{G}_{M+N+1}^{\mathrm{T}})^{\mathrm{T}}$.

[7] The non-zero components $\mathscr{Y}_{ij}$ of the matrix $\mathscr{Y}$ can be calculated in a simple manner from the components of the vectors $Y(n) = (y_1(n), \dots, y_q(n))^{\mathrm{T}}$ using

$$\mathscr{Y}_{[i+1+jq(N+1)](j+1)} = y_{[i \bmod q+1]}(n - i\%q),$$

$$\mathscr{Y}_{[i+1+(j+1)q(N+1)](j+1)} = -y_{[i \bmod q+1]}(n + 1 - i\%q),$$

where mod is modulo, $i\%q$ is the quotient of i divided by q, $0 \leqslant i < q(N+1)$ and $0 \leqslant j < (N+M)$.

The minimization of the cost function in Eq. (12) does not yield the Moore–Penrose generalized inverse (pseudoinverse) of the Sylvester matrix $\boldsymbol{T}_N(\mathscr{H})$, but a $(M+N+1)p \times q(N+1)$ matrix $\boldsymbol{G}$ which satisfies

$$\boldsymbol{G}\boldsymbol{T}_N(\mathscr{H}) = \begin{pmatrix} \boldsymbol{A} & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \boldsymbol{A} & \mathbf{0} & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \mathbf{0} & \ddots & \ddots & \boldsymbol{A} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \dots & \mathbf{0} & \boldsymbol{A} \end{pmatrix}, \tag{13}$$

where $\boldsymbol{A}$ is an arbitrary $p \times p$ matrix (see Appendix A). Using Eqs. (5) and (13), we find that

$$\boldsymbol{G}Y_N(n) = \begin{pmatrix} \boldsymbol{A}S(n) \\ \vdots \\ \boldsymbol{A}S(n-M-N) \end{pmatrix}. \tag{14}$$

So as the algorithm converges, the estimated signals become an instantaneous mixing of the sources (according to matrix $\boldsymbol{A}$).

Finally, to avoid the spurious solution $\boldsymbol{G} = \mathbf{0}$ and force the matrix $\boldsymbol{A}$ to be an invertible matrix, we propose the minimization subject to the constraint

$$\boldsymbol{G}_1 \boldsymbol{R}_Y(n) \boldsymbol{G}_1^{\mathrm{T}} = \boldsymbol{I}_p, \tag{15}$$

where $\boldsymbol{G}_1$ is the first block row $p \times q(N+1)$ of $\boldsymbol{G}$, $\boldsymbol{R}_Y(n) = EY_N(n)Y_N(n)^{\mathrm{T}}$ is the covariance matrix of $Y_N(n)$ and $\boldsymbol{I}_p$ is a $p \times p$ identity matrix. If the above constraint is satisfied and $\boldsymbol{G}_1$ is such that $\boldsymbol{G}_1 \boldsymbol{Y}_N(n) = \boldsymbol{A}S(n)$, then

$$\boldsymbol{G}_1 \boldsymbol{R}_Y(n) \boldsymbol{G}_1^{\mathrm{T}} = \boldsymbol{A}\boldsymbol{R}_S(n)\boldsymbol{A}^{\mathrm{T}} = \boldsymbol{I}_p, \tag{16}$$

where $\boldsymbol{R}_S(n) = ES(n)S(n)^{\mathrm{T}}$ is the source covariance matrix. $\boldsymbol{R}_S(n)$ is a full rank diagonal matrix as a result of the statistical independence of the p sources from each other. When Eq. (16) is satisfied, matrix $\boldsymbol{A}$ becomes invertible. So, separation of the residual instantaneous mixture becomes possible using any algorithm for the separation of an instantaneous mixture (see Appendix B).

In our simulation, the residual instantaneous mixture is separated according to Ref. [34]. In that paper, the blind separation of an instantaneous mixture is done using a Levenberg–Marquardt method to minimize a cost function based on the fourth-order cross-cumulant.

4. Algorithm

In order to experimentally improve the performance of our algorithm, we attempted to minimize the cost function in Eq. (12) using a conjugate gradient algorithm [7]. The algorithm proposed by Chen et al. [7] can minimize a cost function $f(V)$ with respect to a vector V. In theory, this algorithm can converge in a number of iterations which is less than the dimension of V.

In our case, the cost function (12) must be minimized with respect to a $p \times q(N+1)(M+N+1)$ matrix $\mathscr{G}$. It was shown, in our previous study [23], that such an operation can be performed using the Col[8] operator.

[8] The operator Col accords to a $m \times n$ matrix $G = (g_{ij})$ a mn vector $V = (v_i)$ such $v_{i+(j-1)m} = g_{ij}$.

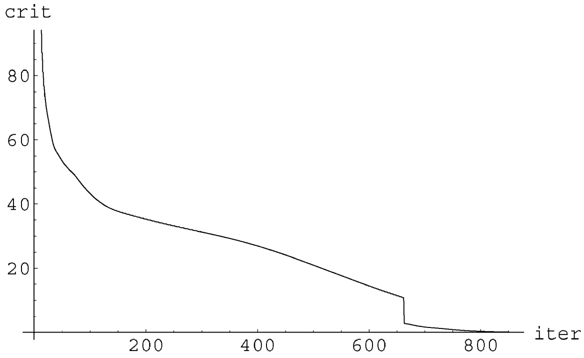

Fig. 2. The convergence of the sub-space criterion.

Unfortunately, this approach has many experimental problems due to the minimization of a cost functions containing huge matrices, see Ref. [23]. Consequently, we suggest here, that the cost function in Eq. (12) should be decomposed into p cost functions, where each one is dependent only on one row of $\mathscr{G}$ (see Appendix C). We can then easily apply the conjugate gradient algorithm to minimize our criterion.

Finally, the constraint in Eq. (15) can be satisfied easily by a simple Cholesky decomposition, then $\boldsymbol{G}_1^*$ can be normalized by $\boldsymbol{G}_1^* = (\boldsymbol{G}_1\boldsymbol{R}_Y(n)\boldsymbol{G}_1^{\mathrm{T}})^{-1/2}\boldsymbol{G}_1$ at each iteration. In addition, the source separation of the instantaneous residual mixture is achieved according to the method proposed in Ref. [30] and the estimates of the different statistics are achieved according to Ref. [31].

5. Experimental results

Many experimental studies show that for two stationary sources, the convergence of the subspace criterion (12) is attained with less than 1000 iterations. In the experiment shown in this section, the convergence was attained with 800 iterations (see Fig. 2).

In that experiment, four sensors $q = 4$ and two stationary sources $p = 2$ are used:

- The first source is a signal with a uniform probability density function (pdf).
- The second signal is the output of an moving average (MA) filter $h(z) = 1 + 0.5z^{-1} - 0.4z^{-2} + 0.2z^{-3}$, which has a signal with uniform pdf as the input.

The channel effect $\mathscr{H}(z)$ is considered as a 4×2 matrix of finite-duration impulse response (FIR) filters of fourth degree ($M = 4$):

$$\mathscr{H}(z) = \begin{pmatrix} -1 - 2z^{-1} + z^{-2} + 1.5z^{-3} + z^{-4} & z^{-1} + z^{-2} + 2z^{-3} + 1.5z^{-4} \\ 2 - 4z^{-1} + 4z^{-2} & 1 - 2z^{-1} + 1.5z^{-2} + z^{-3} + 0.5z^{-4} \\ -1 - z^{-1} + 0.4z^{-2} + 3z^{-3} - z^{-4} & 3 - 2z^{-2} + 2z^{-3} + z^{-4} \\ -2 + z^{-2} + 4z^{-3} - 1.5z^{-4} & 1 + 2z^{-1} - 2.5z^{-2} - z^{-3} + 0.4z^{-4} \end{pmatrix}. \tag{17}$$

Fig. 3 shows that the objective of first step of the algorithm was achieved, with $G.\boldsymbol{T}_N(\mathscr{H})$ being a block diagonal matrix (where $\boldsymbol{A}$ is a 2×2 matrix, see Eq. (13)). Therefore, the two ($p = 2$) output signals $z_i(n)$ are given by $Z(n) = (z_1(n), z_2(n))^{\mathrm{T}} = \boldsymbol{A}S(n)$, and the separation of the instantaneous residual mixture is achieved using the instantaneous algorithm [30]. The convergence speed of this step is shown in Fig. 4.

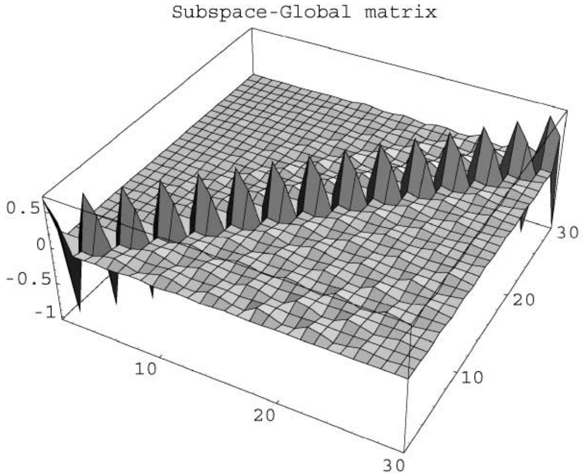

Fig. 3. Performance results: $G.\boldsymbol{T}_N(\mathscr{H})$ should be a block diagonal matrix.

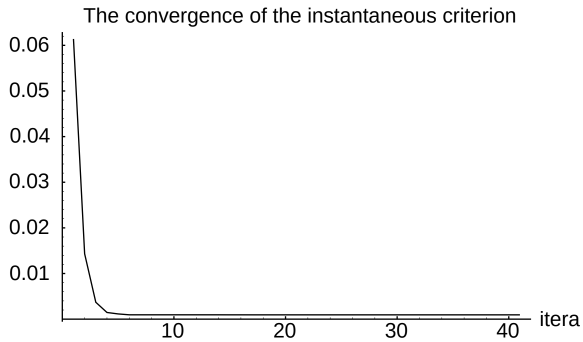

Fig. 4. Performances of the instantaneous residual mixture separation.

Finally Fig. 5 shows the behavior of our algorithm and its performance.[9] In the figure, we plot the various signals in their own planes. In Fig. 5, it should be noted that the sources $s_1(n)$ and $s_2(n)$ are statistically independent as are the estimated signals $x_1(n)$ and $x_2(n)$ (because we obtained a rectangular shape[10]). In addition, Fig. 5(c) shows that the output signals may be obtained by mixing independent sources with the help of an instantaneous mixture. Finally, the mixing signals are given in Fig. 5(b) and the estimated sources in Fig. 5(d).

Even if the convergence of this algorithm is attained within a small number of iterations (in general, less than 1000 iterations are needed), the convergence time is relatively important due to large matrices in the cost function. In effect, we are working toward improving the algorithm convergence, so we can separate numerous sources within a short time.

6. Conclusion

In this paper, we present a new subspace adaptive algorithm to estimate the sources, using the statistics of observed signals issued from a convolutive mixing of the sources.

The main idea behind this algorithm is the use of a subspace approach to obtain the inverse of the Sylvester matrix corresponding to the channel effects. In other words, by minimizing a second-order statistics criterion, we can simplify the original problem by transforming the mixture from a convolutive one to an instantaneous one. The separation of the residual instantaneous mixture can be done using any instantaneous mixture

[9] Concerning the sensitivity of the algorithm to noise, we found, in a recent study of a similar approach [35], that the performance of the algorithm is satisfactory for reasonable RSB (around 20 dB).

[10] For more information concerning the relationship between the distribution of signals and the geometrical form plotted by these signals in their own plane, see Ref. [37].

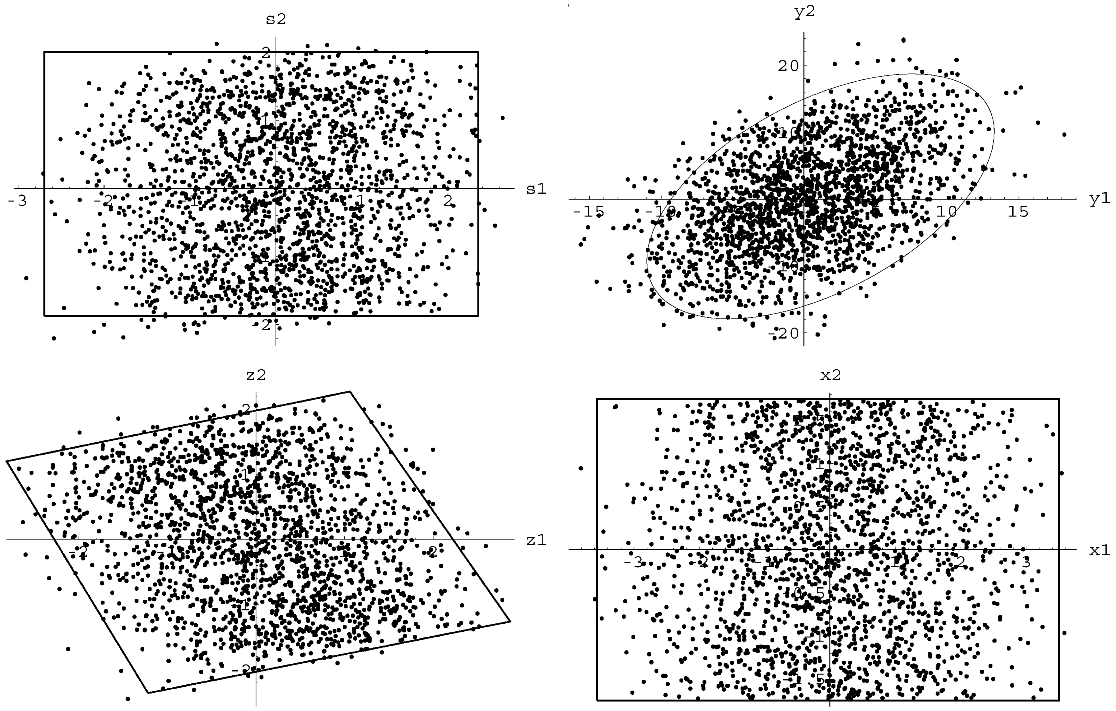

Fig. 5. Experimental results: (a) sources signals $s_1 - s_2$; (b) mixing signals $y_1 - y_2$; (c) first step of the sub-space algorithm $z_1 - z_2$; (d) estimated signals $x_1 - x_2$.

algorithm, typically based on fourth-order statistics. Consequently, we find that most of the inverse channel parameters can be estimated using only second-order statistics.[11]

To improve the convergence speed and the performance of the algorithm proposed in this paper, the minimization of the proposed second-order statistics criterion was achieved using a conjugate gradient method.

Finally, the experimental study shows that for stationary signals, the algorithm convergence is performed in less than a 1000 iterations and satisfactory experimental results were obtained (the cross-talk is about − 22 dB). Unfortunately, we did not obtain similar results for nonstationary signals such as speech signals. We are aiming to improve the algorithm by modifying the criterion and the constraint, or using a preprocessing algorithm [35]. Therefore, it can yield satisfactory results even if the sources are strongly non-stationary signals such as speech signals.

Acknowledgements

The author is grateful to Prof. Christian Jutten (UJF, France) and to Prof. Philippe Loubaton (Univ. de la Marne la Vallée, France) for discussions and comments. He would like also to thank the anonymous referees and the editor for their suggestions and comments.

[11] The mixing matrix $\boldsymbol{A}$ of the residual mixture is $p \times p$, therefore the high-order statistics are used to estimate only p^2 coefficients instead of pqM coefficients of $\mathscr{H}(z)$ (for example, when $p = 2$, $q = 4$ and $M = 10$, there are 80 coefficients in $\mathscr{H}(z)$ and only four coefficients in $\boldsymbol{A}$).

Appendix A. Solution: type and uniqueness

In this section, we will show that a matrix $\boldsymbol{G}$ satisfies (11) if

$$\boldsymbol{G}\boldsymbol{T}_N(\mathscr{H}) = \begin{pmatrix} \boldsymbol{A} & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \boldsymbol{A} & \mathbf{0} & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \mathbf{0} & \ddots & \ddots & \boldsymbol{A} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \dots & \mathbf{0} & \boldsymbol{A} \end{pmatrix}, \tag{A.1}$$

where $\boldsymbol{A}$ is an arbitrary $p \times p$ matrix. In fact, it is easy to prove, from Eq. (11), that

$$\boldsymbol{G}_i Y_N(n) = \boldsymbol{G}_{i+1} Y_N(n+1). \tag{A.2}$$

Using Eq. (5), Eq. (A.2) can be rewritten as

$$\boldsymbol{G}_i \boldsymbol{T}_N(\mathscr{H}) S_{M+N}(n) = \boldsymbol{G}_{i+1} \boldsymbol{T}_N(\mathscr{H}) S_{M+N}(n+1), \tag{A.3}$$

$\forall 1 \leqslant i \leqslant M+N$. Using this Eq. (A.3), we can prove that

$$[I_{(M+N)p} \quad \mathbf{0}_p]\boldsymbol{G}\boldsymbol{T}_N(\mathscr{H}) S_{M+N}(n) = [\mathbf{0}_p \quad I_{(M+N)p}]\boldsymbol{G}\boldsymbol{T}_N(\mathscr{H}) S_{M+N}(n+1), \tag{A.4}$$

where $I_{(M+N)p}$ is the $(M+N)p \times (M+N)p$ identity matrix and $\mathbf{0}_p$ is a $(M+N)p \times p$ zero matrix. Let $\mathscr{A}$ denote a $(M+N+1)p \times (M+N+1)p$ matrix, such that $\boldsymbol{G}\boldsymbol{T}_N(\mathscr{H}) = \mathscr{A}$. Using definition (4) and Eq. (A.4), we can write

$$[\mathbf{0}_p \quad [I_{(M+N)p} \quad \mathbf{0}_p]\mathscr{A}] S_{M+N+1}(n+1) = [[\mathbf{0}_p \quad I_{(M+N)p}]\mathscr{A} \quad \mathbf{0}_p] S_{M+N+1}(n+1). \tag{A.5}$$

Let $\mathscr{B}$ denote the $(M+N)p \times (M+N+2)p$ matrix defined by

$$\mathscr{B} = [\mathbf{0}_p \; [I_{(M+N)p} \; \mathbf{0}_p]\mathscr{A}] - [[\mathbf{0}_p \; I_{(M+N)p}] \; \mathscr{A} \; \mathbf{0}_p].$$

Additionally, let us denote by $\mathscr{V}_n$ the $(M+N+2)p$-dimensional vector defined by $\mathscr{V}_n = S_{M+N+1}(n)$. Eq. (A.5) can then be written as

$$\mathscr{B}\mathscr{V}_{n+1} = 0. \tag{A.5a}$$

From Eq. (A.5a), one can conclude that

$$\mathscr{V}_{n+1} \in \mathrm{Null}\{\mathscr{B}\}, \tag{A.5b}$$

where Null is the null space of $\mathscr{B}$. Assuming that the sources are persistently exciting such that one can obtain $(M+N+2)p$ linearly independent vectors $\mathscr{V}_i$, $i \in \{\xi_1, \dots, \xi_{(M+N+2)p}\}$ and ξ_j are integers such as $\xi_1 < \xi_2 < \cdots < \xi_{(M+N+2)p}$. In this case, using Eq. (A.5b) and the fact that Eq. (A.5) should be satisfied for every n, one can write that

$$\dim\{\mathrm{Null}\{\mathscr{B}\}\} = (M+N+2)p. \tag{A.5c}$$

On the other hand, it is known [18] that

$$\dim\{\mathrm{Null}\{\mathscr{B}\}\} + \mathrm{Rank}\{\mathscr{B}\} = (M+N+2)p. \tag{A.5d}$$

Using Eqs. (A.5c) and (A.5d), one can conclude that $\mathrm{Rank}\{\mathscr{B}\} = 0$ or that $\mathscr{B} = \mathbf{0}$. Therefore, one can write

$$[\mathbf{0}_p \quad [I_{(M+N)p} \quad \mathbf{0}_p]\mathscr{A}] = [[\mathbf{0}_p \quad I_{(M+N)p}]\mathscr{A} \quad \mathbf{0}_p] \tag{A.6}$$

Finally, we can represent matrix $\mathscr{A}$ in different ways:

$$\mathscr{A} = (\boldsymbol{A}_{ij}) = \begin{bmatrix} \boldsymbol{A}_1 \\ \boldsymbol{A}_2 \end{bmatrix} = \begin{bmatrix} \boldsymbol{A}_3 \\ \boldsymbol{A}_4 \end{bmatrix}, \tag{A.7}$$

where $\boldsymbol{A}_{ij}$ is a $p \times p$ matrix, $\boldsymbol{A}_1$ and $\boldsymbol{A}_4$ are $p \times (M+N+1)p$ matrix and $\boldsymbol{A}_2$ and $\boldsymbol{A}_3$ are the $(M+N)p \times (M+N+1)p$ matrix. Using Eqs. (A.7) and (A.6), it is easy to show that

$$[\boldsymbol{0}_p \quad \boldsymbol{A}_3] = [\boldsymbol{A}_2 \quad \boldsymbol{0}_p] \Rightarrow \begin{cases} \boldsymbol{A}_{i1} = \boldsymbol{0} & \forall 2 \leqslant i \leqslant M+N+1, \\ \boldsymbol{A}_{i(M+N+1)} = \boldsymbol{0} & \forall 1 \leqslant i \leqslant M+N, \\ \boldsymbol{A}_{ij} = \boldsymbol{A}_{(i-1)(j-1)} & \forall 1 \leqslant i \leqslant M+N \text{ and } 1 \leqslant j \leqslant M+N. \end{cases} \tag{A.8}$$

From Eq. (A.8), Eq. (A.1) is easily derived.

Appendix B. Consistent of the criterion and the constraint

In this section, we answer the question: Are Eqs. (12) and (15) consistent? From Appendix A, we know that the solution of Eq. (12) belongs to a set of matrices Ω such that

$$\text{if } \boldsymbol{G} \in \Omega \Leftrightarrow \boldsymbol{G}\boldsymbol{T}_N(\mathscr{H}) = \begin{pmatrix} \boldsymbol{A} & \boldsymbol{0} & \boldsymbol{0} & \cdots & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{A} & \boldsymbol{0} & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \boldsymbol{0} & \ddots & \ddots & \boldsymbol{A} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0} & \dots & \boldsymbol{0} & \boldsymbol{A} \end{pmatrix}, \tag{B.1}$$

where $\boldsymbol{A}$ is a $p \times p$ matrix. On the other hand, the output of the subspace part $Z(n)$ is obtained by

$$Z(n) = [\boldsymbol{I}_p, \boldsymbol{0}, \dots, \boldsymbol{0}]\boldsymbol{G}Y_N(n). \tag{B.2}$$

In this case, one can rewrite constraint (16) as

$$\boldsymbol{R}_Z = EZ(n)Z^{\mathrm{T}}(n) = \boldsymbol{I}_p. \tag{B.3}$$

Using the two Eqs. (B.1) and (B.3), one can prove that matrix $\boldsymbol{A}$ belongs to a set of matrices $\mho$:

$$\boldsymbol{A} \in \mho \Leftrightarrow \boldsymbol{A}\boldsymbol{R}_S(n)\boldsymbol{A}^{\mathrm{T}} = \boldsymbol{I}_p. \tag{B.4}$$

Let $\boldsymbol{K}$ denote a square root of $\boldsymbol{R}_S(n)$ ($\boldsymbol{K}$ can be obtained by different methods, such as Cholesky's method [15]). Let $\boldsymbol{\Psi} = \boldsymbol{K}^{-1}$; the matrix $\boldsymbol{K}$ is a full rank matrix because $\boldsymbol{R}_S(n)$ is a full rank matrix. Now, one can rewrite

$$\boldsymbol{A} \in \mho \Leftrightarrow \exists \boldsymbol{\Theta} \,|\, \boldsymbol{A} = \boldsymbol{\Theta}\boldsymbol{\Psi}, \tag{B.5}$$

here $\boldsymbol{\Theta}$ is any $p \times p$ orthogonal matrix. Therefore, $\boldsymbol{A}$ can be obtained up to an orthogonal matrix and one needs another stage based on high-order statistics to achieve the separation.

Appendix C. The algorithm for two sources

To explain our idea, let us consider the simple case of two sources $p = 2$. Let us denote the ith row of $\mathscr{G}$ by $\mathscr{G}_i$. Now, constraint (15) can be rewritten as

$$\begin{pmatrix} \mathscr{G}_1 \\ \mathscr{G}_2 \end{pmatrix} \begin{pmatrix} \boldsymbol{R}_Y(n) & \mathbf{0}_{q(N+1)} \\ \mathbf{0}^{\mathrm{T}}_{q(N+1)} & \mathbf{0}_{q(M+N)(N+1)} \end{pmatrix} (\mathscr{G}_1^{\mathrm{T}} \quad \mathscr{G}_2^{\mathrm{T}}) = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \tag{C.1}$$

where $\mathbf{0}_{q(N+1)}$ is a $q(N+1) \times q(M+N)(N+1)$ zero matrix and $\mathbf{0}_{q(M+N)(N+1)}$ is a $q(M+N)(N+1) \times q(M+N)(N+1)$ zero matrix. Now, it is easy to show that the algorithm can be divided into two steps:

- The first step involves of the estimation of $\mathscr{G}_1$. Therefore, the cost function (12) and the constraint (C.1) become

$$\begin{cases} \min_{\mathscr{G}_1} \mathscr{G}_1 \sum_{n=n_0}^{n_1} \mathscr{Y}(n)\mathscr{Y}^{\mathrm{T}}(n)\mathscr{G}_1^{\mathrm{T}} \\ \text{with respect to } \mathscr{G}_1 \begin{pmatrix} \boldsymbol{R}_Y(n) & \mathbf{0}_{q(N+1)} \\ \mathbf{0}^{\mathrm{T}}_{q(N+1)} & \mathbf{0}_{q(M+N)(N+1)} \end{pmatrix} \mathscr{G}_1^{\mathrm{T}} = 1. \end{cases} \tag{C.2}$$

- The second step involves the estimation of $\mathscr{G}_2$. In this case, cost function (12) and constraint (C.1) become

$$\begin{cases} \min_{\mathscr{G}_2} \mathscr{G}_2 \sum_{n=n_0}^{n_1} \mathscr{Y}(n)\mathscr{Y}^{\mathrm{T}}(n)\mathscr{G}_2^{\mathrm{T}} \\ \text{with respect to } \mathscr{G}_2 \begin{pmatrix} \boldsymbol{R}_Y(n) & \mathbf{0}_{q(N+1)} \\ \mathbf{0}^{\mathrm{T}}_{q(N+1)} & \mathbf{0}_{q(M+N)(N+1)} \end{pmatrix} \mathscr{G}_2^{\mathrm{T}} = 1 \\ \text{and } \mathscr{G}_1 \begin{pmatrix} \boldsymbol{R}_Y(n) & \mathbf{0}_{q(N+1)} \\ \mathbf{0}^{\mathrm{T}}_{q(N+1)} & \mathbf{0}_{q(M+N)(N+1)} \end{pmatrix} \mathscr{G}_2^{\mathrm{T}} = 0. \end{cases} \tag{C.3}$$

Eq. (C.3) can be derived as:

$$\begin{cases} \min_{\mathscr{G}_2} \mathscr{G}_2 \Bigg[\sum_{n=n_0}^{n_1} \mathscr{Y}(n)\mathscr{Y}^{\mathrm{T}}(n) + \\ \begin{pmatrix} \boldsymbol{R}_Y(n) & \mathbf{0}_{q(N+1)} \\ \mathbf{0}^{\mathrm{T}}_{q(N+1)} & \mathbf{0}_{q(M+N)(N+1)} \end{pmatrix} \mathscr{G}_1^{\mathrm{T}} \mathscr{G}_1 \begin{pmatrix} \boldsymbol{R}_Y(n) & \mathbf{0}_{q(N+1)} \\ \mathbf{0}^{\mathrm{T}}_{q(N+1)} & \mathbf{0}_{q(M+N)(N+1)} \end{pmatrix} \Bigg] \mathscr{G}_2^{\mathrm{T}} \\ \text{with respect to } \mathscr{G}_2 \begin{pmatrix} \boldsymbol{R}_Y(n) & \mathbf{0}_{q(N+1)} \\ \mathbf{0}^{\mathrm{T}}_{q(N+1)} & \mathbf{0}_{q(M+N)(N+1)} \end{pmatrix} \mathscr{G}_2^{\mathrm{T}} = 1. \end{cases} \tag{C.4}$$

We can easily apply the conjugate gradient algorithm to minimize our criterion, in the two cases given in Eqs. (C.2) and (C.4). In addition, constraints (C.2) and (C.4) can be satisfied easily by a simple Cholesky decomposition.[12]

[12] By example, $\mathscr{G}_1$ can be normalized by $\mathscr{G}_1^* = (\mathscr{G}_1 \mathscr{R}_Y(n)\mathscr{G}_1^{\mathrm{T}})^{-1/2}\mathscr{G}_1$ at each iteration.

References

[1] K. Abed Meraim, J.F. Cardoso, A. Gorokhov, P. Loubaton, E. Moulines, On subspace methods for blind identification of single-input multiple-output fir systems, IEEE Trans. Signal Process. 45 (1) (January 1997) 42–55.

[2] K. Abed Meraim, P. Loubaton, E. Moulines, A subspace algorithm for certain blind identification problems, IEEE Trans. Inform. Theory 43 (3) (March 1997) 499–511.

[3] S.I. Amari, J.F. Cardoso, Blind source separation-semiparametric statistical approach, IEEE Trans. Signal Process. 45 (11) (November 1997) 2692–2700.

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

[5] J.F. Cardoso, P. Comon, Tensor-based independent component analysis, Signal Processing 5 (1990) 673–676.

[6] J.F. Cardoso, B. Laheld, Equivariant adaptive source separation, IEEE Trans. Signal Process. 44 (12) (December 1996) 3017–3030.

[7] H. Chen, T.K. Sarkar, S.A. Dianat, J.D. Brule, Adaptive spectral estimation by the conjugate gradient method, IEEE Trans. Acoust. Speech Signal Process. ASSP-34 (2) (April 1986) 272–284.

[8] S. Dégerine, R. Malki, Second order blind separation of sources based on canonical partial innovations, IEEE Trans. Signal Process. 48 (3) (March 2000) 629–641.

[9] L. De Lathauwer, D. Callaerts, B. De Moor, J. Vandewalle, Separation of wide band sources, HOS 95, Girona, Spain, 12–14 June 1995, pp. 134–138.

[10] N. Delfosse, P. Loubaton, Adaptive blind separation of convolutive mixtures, Proceedings of ICASSP, Atlanta, GA, May 1996, pp. 2940–2943.

[11] G. Desodt, D. Muller, Complex independent components analysis applied to the separation of radar signals, in: L. Torres, E. Masgrau, M.A. Lagunas (Eds.), Signal Processing V, Theories and Applications, Barcelona, Espain, Elsevier, Amsterdam, 1994, pp. 665–668.

[12] G. D'urso, L. Cai, Sources separation method applied to reactor monitoring, Proceedings of Workshop Athos Working Group, Girona, Spain, June 1995.

[13] D. Gesbert, P. Duhamel, 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, W.A. Sandham (Eds.), Signal Processing VII, Theories and Applications, Edinburgh, Scotland, Elsevier, Amsterdam, September 1994, pp. 712–715.

[14] D. Gesbert, P. Duhamel, S. Mayrargue, On-line blind multichannel equalization based on mutually referenced filters, IEEE Trans. Signal Process. 45 (9) (September 1997) 2307–2317.

[15] G.H. Golub, C.F. Van Loan, Matrix Computations, The Johns Hopkins Press, London, 1984.

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

[17] J. Hérault, B. Ans, Réseaux de neurones à synapses modifiables: Décodage de messages sensoriels composites par une apprentissage non supervisé et permanent, C. R. Acad. Sci. Paris, sér. III (1984) 525–528.

[18] R.A. Horn, Ch.R. Johnson, Matrix Analysis, Cambridge University Press, Cambridge, 1985.

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

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

[21] T. Kailath, Linear Systems, Prentice-Hall, Englewood Cliffs, NJ, 1980.

[22] U. Lindgren, H. Broman, Source separation using a criterion based on second-order statistics, IEEE Trans. Signal Process. 46 (7) (July 1998) 1837–1850.

[23] A. Mansour, Contributions à la séparation de sources, Ph.D. Thesis, INPG Grenoble, 12 January 1997.

[24] A. Mansour, C. Jutten, Fourth order criteria for blind separation of sources, IEEE Trans. Signal Process. 43 (8) (August 1995) 2022–2025.

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

[26] A. Mansour, C. Jutten, A direct solution for blind separation of sources, IEEE Trans. Signal Process. 44 (3) (March 1996) 746–748.

[27] A. Mansour, C. Jutten, P. Loubaton, Subspace method for blind separation of sources and for a convolutive mixture model, in: Signal Processing VIII, Theories and Applications, Triest, Italy, Elsevier, Amsterdam, September 1996, pp. 2081–2084.

[28] A. Mansour, C. Jutten, P. Loubaton, Robustesse des hypothèses dans une méthode sousespace pour la séparation de sources, Actes du XVIème colloque GRETSI, Grenoble, France, September 1997, pp. 111–114.

[29] A. Mansour, C. Jutten, P. Loubaton, An adaptive subspace algorithm for blind separation of independent sources in convolutive mixture, IEEE Trans. Signal Process. 48 (2) (February 2000) 583–586.

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

[31] A. Mansour, A. Kardec Barros, N. Ohnishi, Comparison among three estimators for high order statistics, Fifth International Conference on Neural Information Processing (ICONIP'98), Kitakyushu, Japan, 21–23 October 1998, pp. 899–902.
[32] A. Mansour, A. Kardec Barros, N. Ohnishi, Blind separation of sources: methods, assumptions and applications, IEICE Trans. Fund. Electron. Commun. Comput. Sci. E83-A(8) (2000) 1498–1512 (Special Section on Digital Signal Processing in IEICE EA).
[33] A. Mansour, N. Ohnishi, A blind separation algorithm based on subspace approach, IEEE-EURASIP Workshop on Nonlinear Signal and Image Processing (NSIP'99), Antalya, Turkey, 20–23 June 1999, pp. 268–272.
[34] A. Mansour, N. Ohnishi, Multichannel blind separation of sources algorithm based on cross-cumulant and the Levenberg–Marquardt method, IEEE Trans. Signal Process. 47 (11) (November 1999) 3172–3175.
[35] A. Mansour, N. Ohnishi, A batch subspace ICA algorithm, 10th IEEE Signal Processing Workshop on Statistical Signal and Array Processing, Pocono Manor Inn, PA, USA, 14–16 August 2000, pp. 63–67.
[36] L. Nguyen Thi, C. Jutten, Blind sources separation for convolutive mixtures, Signal Processing 45 (2) (1995) 209–229.
[37] G. Puntonet, A. Mansour, C. Jutten, Geometrical algorithm for blind separation of sources, Actes du XVème colloque GRETSI, Juan-Les-Pins, France, 18–21 September 1995, pp. 273–276.
[38] N. Thirion, J. Mars, J.L. Boelle, Separation of seismic signals: a new concept based on a blind algorithm, in: Signal Processing VIII, Theories and Applications, Triest, Italy, Elsevier, Amsterdam, September 1996, pp. 85–88.
[39] S. Van Gerven, D. Van Compernolle, Signal separation by symmetric adaptive decorrelation: stability, convergence, and uniqueness, IEEE Trans. Signal Process. 43 (7) (July 1995) 1602–1612.
[40] S. Van Gerven, D. Van Compernolle, L. Nguyen Thi, C. Jutten, Blind separation of sources: a comparative study of a 2nd and a 4th order solution, in: M.J.J. Holt, C.F.N. Cowan, P.M. Grant, W.A. Sandham (Eds.), Signal Processing VII, Theories and Applications, Edinburgh, Scotland, Elsevier, Amsterdam, September 1994, pp. 1153–1156.

A. Mansour received his Electronic-Electrical Engineering Diploma in 1992 from the Lebanese University (Tripoli, Lebanon), and his M.Sc. and the Ph.D. degrees in Signal, Image and Speech Processing from the Institut National Polytechnique de Grenoble – INPG (Grenoble, France) in August 1993 and January 1997, respectively. From January 1997 to July 1997, he held a post-doc position at Laboratoire de Traitement d'Images et Reconnaissance de Forme at the INPG, Grenoble, France. Since August 1997, he has been a Research Scientist at the Bio-Mimetic Control Research Center (BMC) at the Institut of Physical and Chemical Research (RIKEN), Nagoya, Japan. His research interests are in the areas of blind separation of sources, high-order statistics, signal processing and robotics. He is the first author of many papers published in international journals, such as IEEE Trans on Signal Processing, IEEE Signal Processing Letters, NeuroComputing, IEICE, Alife & Robotics. He is also the first author of many papers published in the proceedings of various international conferences.