Blind Extraction of Singularly Mixed Source Signals

Yuanqing Li, Jun Wang, *Senior Member, IEEE*, and Jacek M. Zurada, *Fellow, IEEE*

***Abstract*—This paper introduces a novel technique for sequential blind extraction of singularly mixed sources. First, a neural-network model and an adaptive algorithm for single-source blind extraction are introduced. Next, extractability analysis is presented for singular mixing matrix, and two sets of necessary and sufficient extractability conditions are derived. The adaptive algorithm and neural-network model for sequential blind extraction are then presented. The stability of the algorithm is discussed. Simulation results are presented to illustrate the validity of the adaptive algorithm and the stability analysis. The proposed algorithm is suitable for the case of nonsingular mixing matrix as well as for singular mixing matrix.**

***Index Terms*—Adaptive algorithm, blind extraction, extractability, singular matrix, stability.**

I. Introduction

BLIND separation of independent source signals from their mixtures has received considerable attention in recent years. This class of signal processing techniques can be used in many technical areas such as communications, medical signal processing, speech recognition, image restoration, etc. [1].

The objective of blind separation is to recover source signals from their mixture without prior information on the source signals and mixing channel. The mixtures of source signals can be divided into several categories, such as instantaneous mixtures, and dynamical or convolutive mixtures [20]–[25]. Independent component analysis (ICA) can be used for blind separation of instantaneous mixtures, whereas dynamical component analysis (DCA) can deal with convolutive mixtures. The mixing process may also be classified as linear or nonlinear. Although the problems of nonlinear channel equalization and blind source separation in nonlinear mixture have been studied in several papers such as [2] and [3], most technical publications focus on blind separation of linear mixtures of source signals [4]–[11], [17]–[19].

Consider a general linear case of instantaneous mixing of n sources observable at m outputs

$$y(t) = Ax(t) \tag{1}$$

where

$x(t) = [x_1(t), \cdots, x_n(t)]^T$ n-dimensional source signal;
$y(t) = [y_1(t), \cdots, y_m(t)]^T$ m-dimensional mixed signal;
A $m \times n$ constant matrix known as mixing matrix.

In (1), the sources $x(t)$ and the mixing matrix A are unknown and the components of $y(t)$ are observed. The task of blind separation is to recover sources $x_1, \cdots, x_n$ from $y_1, \cdots, y_m$ without knowing A.

An ill-conditioned mixing case has been discussed in [14], [15], where the algorithms of blind extraction have been developed for $n \geq m$. But this approach needs an assumption or preprocessing before blind extraction; that is, y should be prewhitened. The prewhitening processing can be described as follows. When $m \geq n$, set $z = Vy$, where V is an $n \times m$ constant matrix. From the assumptions of independence and unit variance of the components of x, we have

$$\begin{aligned} R_{zz} &= E\left(zz^T\right) = E\left(VAxx^TA^TV^T\right) \\ &= VAE\left(xx^T\right)A^TV^T = VAA^TV^T \end{aligned} \tag{2}$$

where E is the expectation operator. Obviously, A must be of full column rank, otherwise it is impossible to fulfill $R_{zz} = I$, where I is an $n \times n$ identity matrix.

A recent reference [16] presents a neural network and proposes unconstrained extraction and deflation criteria that require neither *a priori* knowledge of source signals nor whitening of mixed signals and can cope with a mixture of signals with positive and negative kurtosis. It is shown that the criteria have no spurious equilibria by proving that all spurious equilibria are unstable. However, as in the previous studies [12], [13], the condition that mixing matrix A is of full column rank is necessary, which means that $n \leq m$. Otherwise, the set $U = \{c\colon c = W^TA,\ W \in R^m\}$ is a subspace of R^n. It is possible that the unstable spurious equilibria in R^n are stable the subspace U, which leads to a spurious solution and the failure of sequential blind extraction when there is no criterion for identifying true solutions and spurious solutions. In [17], a recurrent neural network and its associated learning rule are presented, which can deal with the case in which A is near ill-conditioned or near singular, but has to be nonsingular.

Among numerical aspects of ICA, the case of singular square mixing matrix A deserves special attention and is the focus of this paper. The objective of this paper is to discuss sequential blind extraction of sources when A is $n \times n$ dimensional and singular, thus $y(t)$ is also n dimensional. In this paper, without loss of generality, we assume that the source signals $x_1, \cdots, x_n$ are independent with zero means.

In the ill-conditioned case of singular A, blind extraction is perhaps more effective than blind separation, because the solvability of blind separation for this ill-conditioned case is difficult to obtain. Generally, only one source signal can be obtained by

Manuscript received November 5, 1999; revised May 3, 2000. This workwas supported in part by the Hong Kong Research Grants Council under Grant CUHK4150/97E.

Y. Li is with the Automatic Control Engineering Department, South China University of Technology, Guangzhou, 510641 China.

J. Wang is with the Automation and Computer-Aided Engineering Department, The Chinese University of Hong Kong, Shatin, New Territories, Hong Kong.

J. M. Zurada is with the Electrical and Computer Engineering Department, University of Louisville, Louisville, KY 40292 USA.

Publisher Item Identifier S 1045-9227(00)09842-8.

a single-source blind extraction. By using sequential blind extraction, we can obtain more than one source signals.

In this paper, a neural-network model and its associate learning rule are developed for sequential blind extraction in the ill-conditioned case with singular A. This approach is also suitable for the case in which A is nonsingular. The extractability analysis of the problem is also presented, and two sets of sufficient and necessary conditions are derived.

This paper is organized as follows. The problem statement is given in Section I. A neural-network model and a corresponding adaptive algorithm for single-source blind extraction are presented in Section II. The extractability analysis then follows in Section III. The neural-network model and an adaptive algorithm for sequential blind extraction are developed in Section IV. Section V discusses the stability of the adaptive algorithm for the case of singularly mixed sources. Simulation results for extracting speech signals are presented in Section VI, one of which confirms the validity of the adaptive algorithm, and another reveals the validity of stability analysis in Section V. The concluding remarks in Section VII summarize the results.

II. Single-Source Extraction

In this section, the algorithm for a single-source blind extraction from the mixed sources is introduced.

Define an evaluation function

$$J = \sum_{j=2}^{n} \left[\left(E\left(f\left(z_1\right) g\left(z_j\right)\right)\right)^2 + \left(E\left(f\left(z_j\right) g\left(z_1\right)\right)\right)^2 \right] \quad (3)$$

where $f(s)$ and $g(s)$ are nonlinear odd functions such as s^3 and $\tanh(s)$ [4]; $z_1, \cdots, z_n$ are the outputs of the neural network defined as follows:

$$\begin{cases} z_1(t) = y_1(t) - c_{12}(t)y_2(t) - \cdots - c_{1n}(t)y_n(t), \\ z_2(t) = y_2(t) - c_{21}(t)z_1(t), \\ \vdots \\ z_n(t) = y_n(t) - c_{n1}(t)z_1(t). \end{cases} \quad (4)$$

Obviously, $J = 0$ when z_1 and z_j are statistically independent for $j = 2, \cdots, n$ (i.e., a signal corresponding to z_1 is extracted). The connection weights $c_{12}, \cdots, c_{1n}, c_{21}, \cdots, c_{n1}$ in (3) are adjusted based on the following continuous-time learning rule:

$$\begin{aligned} \frac{dc_{12}(t)}{dt} &= \alpha_{12} f\left(z_1(t)\right) g\left(z_2(t)\right) \\ &\vdots \\ \frac{dc_{1n}(t)}{dt} &= \alpha_{1n} f\left(z_1(t)\right) g\left(z_n(t)\right) \\ \frac{dc_{21}(t)}{dt} &= \alpha_{21} f\left(z_2(t)\right) g\left(z_1(t)\right) \\ &\vdots \\ \frac{dc_{n1}(t)}{dt} &= \alpha_{n1} f\left(z_n(t)\right) g\left(z_1(t)\right) \end{aligned} \quad (5)$$

where $\alpha_{12}, \cdots, \alpha_{n1}$ are constant step sizes.

Model (4) is depicted in Fig. 1, in which y_1 is referred to as the main extracted signal, and z_1 as the main output. How

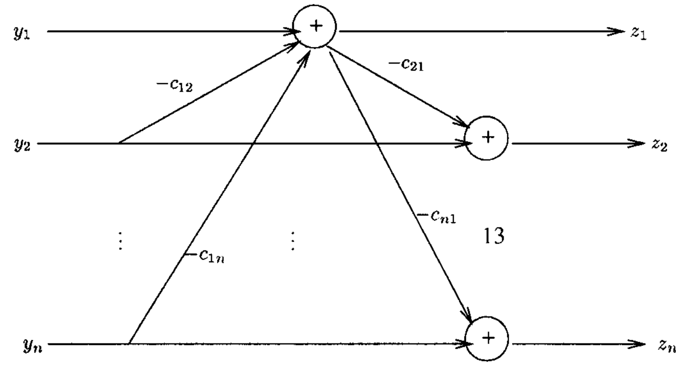


Fig. 1. Single-source blind extraction model.

to determine the main extracted signal will be discussed in the next two sections.

It can be seen from (3) and (4) that if there exist a set of weights $c_{12}, \cdots, c_{1n}, c_{21}, \cdots, c_{n1}$, such that $J = 0$, then the weight vector is an equilibrium of (5).

III. Extractability Analysis

Denote the n row vectors of A as $a_1, a_2, \cdots, a_n$. Since A is singular, the row vectors are linearly dependent and there exists a row vector a_i in A and $(n-1)$ constants $l_2, \cdots, l_n$, such that $a_i = \sum_{j=1, j\neq i}^{n} l_j a_j$.

The following extractability condition will play an important role in blind source extraction.

Extractability Condition: There is a row vector a_i in A which satisfies:

$$a_{ik} = \sum_{m=1, m\neq i}^{n} k_m a_{mk}, \; k = 1, 2, \cdots, n; \; k \neq j \quad (6)$$

$$a_{ij} \neq \sum_{m=1, m\neq i}^{n} k_m a_{mj} \quad (7)$$

where k_m are constants for $m = 1, \cdots, n; \; m \neq i$.

For example, the following mixing matrices satisfy the above extractability condition:

$$\begin{bmatrix} 1 & 1 & 1 \\ 0.5 & 0.5 & 0.5 \\ 0 & 1 & 1 \end{bmatrix}, \; \begin{bmatrix} 0.5 & 2 & 1 & 0.5 \\ 0.1 & 0.5 & 0.5 & 1.0 \\ 0 & 1.5 & 0.5 & -0.5 \end{bmatrix}$$
$$\begin{bmatrix} 1 & 3 & 4 & 5 \\ 3 & 7 & 9 & 11 \\ 0 & 13 & 19 & 7 \\ 5 & 3 & 8 & 31 \end{bmatrix}.$$

Note that the third matrix above is nonsingular. It is not difficult to prove that any nonsingular matrix satisfies the extractability condition. Therefore the subsequent results in this paper are also applicable to nonsingular case.

In the extractability condition, row i in A is singled out for extraction [i.e., y_i is the extracted signal to replace y_1 in (4) and (5)]. For the convenience of discussions, we assume hereafter that a_1 satisfies the extractability condition and y_1 is the main extracted signal. In general, if a_i satisfies the extractability condition, then a signal corresponding to z_i will be extracted.

Lemma: If the extractability condition holds, then there exists a set of weights c_{ij} in (4) such that a single source can be extracted from the mixture of sources.

Proof: Without loss of generality, suppose that a_1 satisfies the extractability condition. Then

$$\begin{bmatrix} 1 & -k_2 & \cdots & -k_n \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix} = \begin{bmatrix} a_{11} - \sum_{j=2}^{n} k_j a_{j1} & 0 & \cdots & 0 \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix}. \tag{8}$$

Denote $b = a_{11} - \sum_{j=2}^{n} k_j a_{j1}$. Obviously $b \neq 0$.

Since

$$\begin{bmatrix} 1 & 0 & \cdots & 0 \\ -\frac{a_{21}}{b} & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ -\frac{a_{n1}}{b} & 0 & \cdots & 1 \end{bmatrix} \begin{bmatrix} b & 0 & \cdots & 0 \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix} = \begin{bmatrix} b & 0 & \cdots & 0 \\ 0 & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & a_{n2} & \cdots & a_{nn} \end{bmatrix} \tag{9}$$

it can be obtained from (1) and (9)

$$\begin{bmatrix} z_(t) \\ z_2(t) \\ \vdots \\ z_n(t) \end{bmatrix} = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ -\frac{a_{21}}{b} & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ -\frac{a_{n1}}{b} & 0 & \cdots & 1 \end{bmatrix} \begin{bmatrix} 1 & -k_2 & \cdots & -k_n \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix} y(t) = \begin{bmatrix} b & 0 & \cdots & 0 \\ 0 & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & a_{n2} & \cdots & a_{nn} \end{bmatrix} \begin{bmatrix} x_1(t) \\ x_2(t) \\ \vdots \\ x_n(t) \end{bmatrix} = \begin{bmatrix} b_{x1}(t) \\ a_{22}x_2(t) + \cdots + a_{2n}x_n(t) \\ \vdots \\ a_{n2}x_2(t) + \cdots + a_{nn}x_n(t) \end{bmatrix}. \tag{10}$$

Denoting the right-hand side of (10) as $z(t)$ yields

$$\begin{aligned} z_1(t) &= y_1(t) - k_2 y_2(t) - \cdots - k_n y_n(t) = b x_1(t) \\ z_2(t) &= y_2(t) - \frac{a_{21}}{b} z_1(t) \\ &\vdots \\ z_n(t) &= y_n(t) - \frac{a_{n1}}{b} z_1(t). \end{aligned} \tag{11}$$

Thus the source signal x_1 can be extracted from the mixed sources up to a coefficient (i.e., $z_1 = b x_1$), and the new mixed sources $z_2, \cdots, z_n$ are only related to $x_2, \cdots, x_n$. □

The blind extraction can be continued to determine $x_2, x_3, \cdots, x_n$ based on the new mixed sources $z_2, \cdots, z_n$. Because the extracted source can be any signal, the order of signals in sequential extraction may change.

In the following, the conditions are derived under which the extractability condition holds.

Theorem 1: The extractability condition holds, if and only if there is an $n \times (n-1)$ submatrix of A denoted as $\overline{A}$, which satisfies $\text{rank}(\overline{A}) < \text{rank}(A)$.

Proof: Sufficiency: Without loss of generality, suppose that

$$\overline{A} = \begin{bmatrix} a_{12} & a_{13} & \cdots & a_{1n} \\ a_{22} & a_{23} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n2} & a_{n3} & \cdots & a_{nn} \end{bmatrix}. \tag{12}$$

Denote $\text{rank}(A) = r$, then $\text{rank}(\overline{A}) = r - 1$. Obviously, $\text{rank}(\overline{A}) = r - 1 < n$, thus the n rows of $\overline{A}$ are linearly dependent. Therefore, there exists a row vector of $\overline{A}$ assumed to be equal to $[a_{12}, \cdots, a_{1n}]$ which satisfies

$$[a_{12}, \cdots, a_{1n}] = l_2 [a_{22}, \cdots, a_{2n}] + \cdots + l_n [a_{n2}, \cdots, a_{nn}] \tag{13}$$

where $l_2, \cdots, l_n$ are constants.

Denote $\overline{B}$ as

$$\overline{B} = \begin{bmatrix} a_{22} & a_{23} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n2} & a_{n3} & \cdots & a_{nn} \end{bmatrix}.$$

From (13), $\text{rank}(\overline{B}) = \text{rank}(\overline{A}) = r - 1$.

Consider the linear equation

$$\overline{B}^T \begin{bmatrix} c_2 \\ c_3 \\ \vdots \\ c_n \end{bmatrix} = \begin{bmatrix} a_{12} \\ a_{13} \\ \vdots \\ a_{1n} \end{bmatrix} \tag{14}$$

where $c_2, \cdots, c_n$ are variables. From (13), (14) has a special solution $l_2, \cdots, l_n$.

Consider the homogeneous equation of (14)

$$\overline{B}^T \begin{bmatrix} c_2 \\ c_3 \\ \vdots \\ c_n \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}. \tag{15}$$

There are $n - r$ basic solutions of (15) denoted as

$$\begin{bmatrix} c_{21} \\ c_{31} \\ \vdots \\ c_{n1} \end{bmatrix}, \cdots, \begin{bmatrix} c_{2(n-r)} \\ c_{3(n-r)} \\ \vdots \\ c_{n(n-r)} \end{bmatrix}. \tag{16}$$

Thus the general solution of (14) can be represented by

$$\begin{bmatrix} c_2 \\ c_3 \\ \vdots \\ c_n \end{bmatrix} = \begin{bmatrix} c_{21} \\ c_{31} \\ \vdots \\ c_{n1} \end{bmatrix} t_1 + \cdots + \begin{bmatrix} c_{2(n-r)} \\ c_{3(n-r)} \\ \vdots \\ c_{n(n-r)} \end{bmatrix} t_{n-r} + \begin{bmatrix} l_2 \\ l_3 \\ \vdots \\ l_n \end{bmatrix} \quad (17)$$

where $t_1, \cdots, t_{n-r}$ are arbitrary.

If the extractability condition is not satisfied for a_1, then for any solution $[c_2, \cdots, c_n]$ of (14)

$$[a_{21}, \cdots, a_{n1}] [c_2, \cdots, c_n]^T = a_{11}. \quad (18)$$

Set $t_1, \cdots, t_{n-r} = 0$, then

$$[a_{21}, \cdots, a_{n1}] [l_2, \cdots, l_n]^T = a_{11}. \quad (19)$$

From (18) and (19), it is easy to obtain

$$[a_{21}, \cdots, a_{n1}] \left[[c_{21}, \cdots, c_{n1}]^T t_1 + \cdots + \left[c_{2(n-r)}, \cdots, c_{n(n-r)}\right]^T t_{n-r} \right] = 0. \quad (20)$$

From that $\text{rank}(A) = r$, (19) and (13), we have $\text{rank}(\overset{*}{A}) = r$, where

$$\overset{*}{A} = \begin{bmatrix} a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix}.$$

Without loss of generality, suppose that $\text{rank}(\overline{A}) = r$, where

$$\overline{A} = \begin{bmatrix} a_{21} & a_{22} & \cdots & a_{2(n-1)} \\ a_{31} & a_{32} & \cdots & a_{3(n-1)} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{n(n-1)} \end{bmatrix}.$$

Consider the equation

$$\overline{A}^T \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_{n-1} \end{bmatrix} = 0 \quad (21)$$

where $v_1, \cdots, v_{n-1}$ are variables. Obviously, the number of basic solutions of (21) is $n - 1 - r$. But from (15), (16) and (20), $[c_{21}, \cdots, c_{n1}]^T, \cdots, [c_{2(n-r)}]^T$ all are the $n - r$ basic solutions of (21). A contradiction has occurred, thus the extractability condition is satisfied. The sufficiency is obtained.

Necessity: It is sufficient to prove that for any $n \times (n - 1)$ dimensional submatrix of A denoted as $\tilde{\overline{A}}$, if $\text{rank}(\tilde{\overline{A}}) = \text{rank}(A)$, then the extractability condition is not satisfied.

According to the conditions above, any column of A can be represented as a linear combination of the remaining $n - 1$ columns.

Without loss of generality, consider the $\overline{A}$ defined in (12). Obviously, there exists a nonzero vector denoted as $[d_2, \cdots, d_n]^T$ which satisfies

$$\begin{bmatrix} a_{11} \\ \vdots \\ a_{n1} \end{bmatrix} = d_2 \begin{bmatrix} a_{12} \\ \vdots \\ a_{n2} \end{bmatrix} + \cdots + d_n \begin{bmatrix} a_{1n} \\ \vdots \\ a_{nn} \end{bmatrix}. \quad (22)$$

For any row $[a_{j2}, \cdots, a_{jn}]$ of $\overline{A}$, if there are constants c_k, $k = 1, \cdots, n$, $k \neq j$, such that

$$[a_{j2}, \cdots, a_{jn}] = \sum_{k=1,\, k\neq j}^{n} c_k \left[a_{k2}, \cdots, a_{kn}\right]$$

then from (22)

$$\begin{aligned} &\left[a_{11}, \cdots, a_{(j-1)1}, a_{(j+1)1}, \cdots, a_{n1}\right] \cdot \\ &\left[c_1, \cdots, c_{(j-1)}, c_{(j+1)}, \cdots, c_n\right]^T \\ &\quad = \sum_{k=2}^{n} d_k \left[a_{1k}, \cdots, a_{(j+1)k}, a_{(j+1)k}, \cdots, a_{nk}\right] \\ &\qquad \cdot \left[c_1, \cdots, c_{(j-1)}, c_{(j+1)}, \cdots, c_n\right]^T \\ &\quad = \sum_{k=2}^{n} d_k a_{jk} = a_{j1} \end{aligned}$$

Thus the extractability condition is not satisfied. The necessity is obtained. □

From Lemma 3 and Theorem 1, we have the following corollary.

Corollary 1: By using (4), there is a set of weights c_{ij} such that a single source can be extracted, if and only if there is an $n \times (n - 1)$ dimensional submatrix $\overline{A}$ of A, which satisfies $\text{rank}(\overline{A}) < \text{rank}(A)$.

The mixed signals sometimes can not be fully separated. For example, let

$$A = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 0 & 0 & 0.5 & 0.3 \\ 0 & 0 & 0.7 & 0.1 \\ 0 & 0 & 0.2 & 0.6 \end{bmatrix}.$$

When the model (4) is used, obviously there is a set of weights c_{ij} such that $z_1 = x_1 + x_2$, $z_2 = y_2 = 0.5x_3 + 0.3x_4$, $z_3 = y_3 = 0.7x_3 + 0.1x_4$, $z_4 = y_4 = 0.2x_3 + 0.6x_4$. Thus $J = 0$, and blind partition is obtained, but x_1 and x_2 can not be separated.

The case above is concluded as follows: When the algorithm in (4) and (5) is used, the first output z_1 is the new mixture of some sources which is independent of the other outputs. The analysis on the case above is presented in the following.

For sake of simplicity, suppose that z_1 is the mixture of x_1 and x_2; i.e.,

$$\begin{aligned} z_1 &= y_1 - \sum_{i=2}^{n} c_{1i} y_i \\ &= \left(a_{11} - \sum_{i=2}^{n} c_{1i} a_{i1}\right) x_1 + \left(a_{12} - \sum_{i=2}^{n} c_{1i} a_{i2}\right) x_2 \end{aligned} \quad (23)$$

where $b_1 = (a_{11} - \sum_{i=2}^{n} c_{1i}a_{i1}) \neq 0$, $b_2 = (a_{12} - \sum_{i=2}^{n} c_{ii}a_{i2}) \neq 0$.

$$\begin{aligned} z_j &= y_j - c_{j1}z_1 \\ &= a_{j1}x_1 + a_{j2}x_2 + \sum_{k=3}^{n} a_{jk}x_k - c_{j1}z_1 \\ &= [a_{j1} - c_{j1}b_1]\, x_1 + [a_{j2} - c_{j1}b_2]\, x_2 + \sum_{k=3}^{n} a_{jk}x_k. \end{aligned} \tag{24}$$

Thus if

$$\frac{a_{j1}}{b_1} = \frac{a_{j2}}{b_2}, \qquad j = 2, \cdots, n \tag{25}$$

then the output z_j is independent of z_1.

Theorem 2: By using the blind extraction model (4), there exists a set of weights c_{ij} such that a mixture of p sources can, be extracted and the other outputs are independent of these sources, if and only if there is an $m \times (n-p)$-dimensional submatrix $\overset{\diamond}{A}$ composed by $n-p$ columns of A, which satisfies $\mathrm{rank}(\overset{\diamond}{A}) < \mathrm{rank}(A)$, and the $m \times p$ submatrix denoted as $\hat{A}$, composed by the remaining p columns of A, has rank 1.

Proof: Necessity: Assume that there is a set of weights c_{ij} in (4) such that a mixture of p sources can be extracted, and the other outputs are independent of these sources. Without loss of generality, suppose that $z_1 = b_1x_1 + \cdots + b_px_p$, where $b_1, \cdots, b_p \neq 0$.

Set

$$\begin{aligned} B &= \begin{bmatrix} 1 & 0 & \cdots & 0 \\ -c_{21} & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ -c_{n1} & 0 & \cdots & 1 \end{bmatrix} \begin{bmatrix} 1 & -c_{12} & \cdots & -c_{1n} \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix} \\ &= \begin{bmatrix} 1 & -c_{12} & \cdots & -c_{1n} \\ -c_{21} & 1 + c_{12}c_{21} & \cdots & c_{1n}c_{21} \\ \vdots & \vdots & \ddots & \vdots \\ -c_{n1} & c_{1n}c_{n1} & \cdots & 1 + c_{1n}c_{n1} \end{bmatrix}. \end{aligned}$$

Then

$$BA = \begin{bmatrix} b_1 & \cdots & b_p & 0 & \cdots & 0 \\ 0 & \cdots & 0 & b_{2(p+1)} & \cdots & b_{2n} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & b_{n(p+1)} & \cdots & b_{nn} \end{bmatrix} = \left[B\hat{A} \middle| B\overset{\diamond}{A} \right] \tag{26}$$

where

$$\hat{A} = \begin{bmatrix} a_{11} & \cdots & a_{1p} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{np} \end{bmatrix}, \tag{27}$$

$$\overset{\diamond}{A} = \begin{bmatrix} a_{1(p+1)}, & \cdots, & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n(p+1)} & \cdots & a_{nn} \end{bmatrix}. \tag{28}$$

In view that B is nonsingular and for (26)

$$\begin{aligned} \mathrm{rank}\left(\overset{\diamond}{A}\right) &= \mathrm{rank}\left(B\overset{\diamond}{A}\right) < \mathrm{rank}(BA) = \mathrm{rank}(A), \\ \mathrm{rank}\left(\hat{A}\right) &= \mathrm{rank}\left(B\hat{A}\right) = 1. \end{aligned}$$

The necessity is obtained.

Sufficiency: Without loss of generality, suppose that $\overset{\diamond}{A}$ defined in (28) satisfies $\mathrm{rank}(\overset{\diamond}{A}) < \mathrm{rank}(A)$, $\hat{A}$ defined in (27) satisfies $\mathrm{rank}(\hat{A}) = 1$. Thus

$$\begin{aligned} &[a_{11}, \cdots, a_{n1}]^T \\ &\quad = s_2\, [a_{12}, \cdots, a_{n2}]^T = \cdots = s_p\, [a_{1p}, \cdots, a_{np}]^T \end{aligned} \tag{29}$$

where $s_2, \cdots, s_p \neq 0$ are constants.

Since $\mathrm{rank}(\overset{\diamond}{A}) < \mathrm{rank}(A)$, there are constants $k_2, \cdots, k_n$, such that one of row of $\overset{\diamond}{A}$ supposed to be $[a_{1(p+1)}, \cdots, a_{1n}]$ can be represented as follows:

$$\begin{aligned} &\left[a_{1(p+1)}, \cdots, a_{1n}\right] \\ &\quad = k_2 \left[a_{2(p+1)}, \cdots, a_{2n}\right] + \cdots + k_n \left[a_{n(p+1)}, \cdots, a_{nn}\right] \end{aligned} \tag{30}$$

and

$$b_l = a_{1l} - \sum_{j=2}^{n} k_j a_{jl} \neq 0, \qquad l = 1, \cdots, p. \tag{31}$$

The justification of (31) is similar to that in the proof of Theorem 1.

From (29), it is easy to obtain

$$b_1 = s_2 b_2 = \cdots = s_p b_p. \tag{32}$$

From (29) and (32), we have

$$\frac{a_{j1}}{b_1} = \cdots = \frac{a_{jp}}{b_p};\; j = 1, \cdots, n. \tag{33}$$

By setting $c_{1j} = k_j$, $c_{j1} = a_{j1}/b_1$, $j = 2, \cdots, n$, then a mixture of p sources can be extracted using model (4), and the other outputs are independent of these sources. The sufficiency is thus obtained. □

Theorem 2 implies that if its conditions are satisfied, then there are p sources which can not be separated, and the largest number of source signals which can be extracted by (4) and (5) is $n - r$.

IV. Sequential Extraction

Before the adaptive algorithm is presented, all possible cases will be discussed first.

Obviously, when the neural network and algorithm (4) and (5) are used, $z_1 = 0$ implies that the first row vector a_1 of A can be represented as a linear combination of the other row vectors.

If $z_1 \neq 0$, and $J = 0$, then either z_1 is one source signal x_i, or z_1 is the mixture of several sources which can not be separated according to Theorem 2. Thus the first extraction is finished, and the next extraction should be continued by using $z_2, \cdots, z_n$ as new mixtures to replace y_i.

There exist three cases of sequential blind extraction.

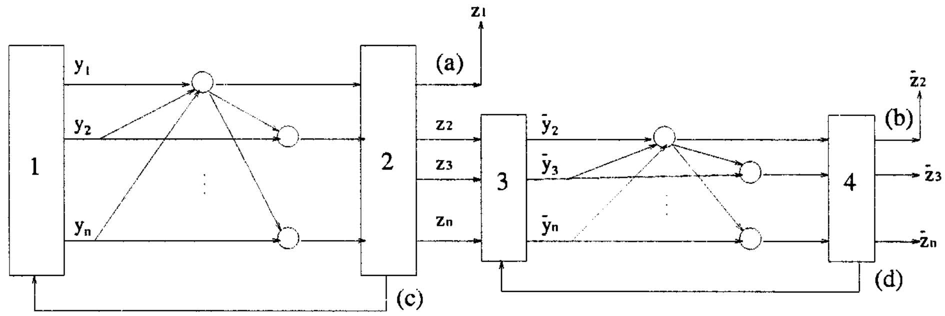


Fig. 2. Diagram of sequential blind extraction.

Case 1: The extractability condition is satisfied for a_1, then a source can be extracted.

Case 2: The conditions in Theorem 2 are satisfied, a mixture of some sources can be extracted, the remaining signals are independent of these sources; i.e., a partition is performed.

If Case 1 or 2 above occurs, then $z_1 \neq 0$ and $J = 0$, the sequential blind extraction continues by using the remaining mixtures.

Case 3: Neither Case 1 nor 2 occurs, then $z_1 = 0$ or $z_1 \neq 0$ and $J \neq 0$, thus any source or a mixture of sources can not be extracted. The main extracted signal should be changed [i.e., y_1 and y_2 are interchanged in (4)], and the single-step blind extraction can be conducted again.

After the mth extraction, if Case 3 always occurs when the main signal has been changed over in the subsequent extraction, then it implies that the extraction algorithm is completed or the remaining sources cannot be separated. In the extracted signals and remaining signals, some are composed of a single source signal x_i, some are mixtures of more source signals.

If the order of extracted signals is neglected, then the final results are the same despite that z_1 is a single x_i or mixture of some x_is.

In order to extract many sources, the sequential extraction should be used. The algorithm is presented below.

Step 1: Initialize $c_{1j}(0), c_{j1}(0) (j = 2, \cdots, n)$, step size α_{ij}, $k = 1$, $l = 0$. Set small positive constants ϵ_0, ϵ_1.

Step 2: If $k \geq n$, stop; otherwise, perform the kth blind extraction

$$\begin{aligned} z_k(t) &= y_k(t) - c_{1(k+1)}(t) y_{k+1}(t) - \cdots - c_{1n}(t) y_n(t) \\ z_{k+l}(t) &= y_{k+1}(t) - c_{(k+1)1}(t) z_k(t) \\ &\vdots \\ z_n(t) &= y_{k+1}(t) - c_{(k+1)1}(t) z_k(t); \end{aligned} \tag{34}$$

$$\begin{aligned} \frac{dc_{1j}(t)}{dt} &= \alpha_{1j} f\left(z_k(t)\right) g\left(z_{k+1}(t)\right), \\ &\vdots \\ \frac{dc_{1n}(t)}{dt} &= \alpha_{1n} f\left(z_k(t)\right) g\left(z_n(t)\right), \\ \frac{dc_{j1}(t)}{dt} &= \alpha_{j1} f\left(z_{k+1}(t)\right) g\left(z_k(t)\right), \\ &\vdots \\ \frac{dc_{n1}(t)}{dt} &= \alpha_{n1} f\left(z_n(t)\right) g\left(z_k(t)\right). \end{aligned} \tag{35}$$

Step 3: Evaluate the results of Step 2 using J.

1) If $J < \epsilon_0$, $\|z_k\| > \epsilon_1$, then z_k is an extracted source signal. Let $y_j = z_j$ $j = k+1, \cdots, n$; $k = k+1$, $l = 0$, go to Step 2.

2) If $J < \epsilon_0$, $\|z_k\| \leq \epsilon_1$, or if $J \geq \epsilon_0$, then set $l = l + 1$, and go to Step 4.

Step 4: If $k+l \leq n$, then exchange the main extracted signal; i.e., set

$$\begin{aligned} y_0 &= y_k \\ y_k &= y_{k+l} \\ y_{k+l} &= y_0 \end{aligned} \tag{36}$$

and go to Step 2; otherwise, stop.

One can decide the values of ϵ_0, ϵ_1 according to their experience. For different problems, these parameters should be set differently. From our simulation experience, the bigger data set of signals to be processed is, the larger ϵ_0 should be set (e.g., ϵ_0 is set as 2.2 in Example 1 in this paper). ϵ_1 can be set very small (e.g., 10^{-3}) to avoid discarding extracted tiny sources of interest. For speech signals, image signals since the intuitionistic results can be used in determining whether the extraction is successful.

A block diagram to illustrate the algorithm is depicted in Fig. 2, where Cases (a) and (b) refer to 1), and Cases (c) and (d) refer to 2) in Step 3.

In the above adaptive algorithm, there is no need for any deflation algorithm. The output of the most recent extraction becomes the input to the current one.

When the mixing matrix is singular, blind extraction can be performed if one of two sets of sufficient and necessary conditions in Corollary 1 and Theorem 2 are satisfied. Otherwise, blind extraction can not be realized by using the adaptive algorithm of this paper. Needless to say, that for the case in which the conditions above are not satisfied, it is difficult to carry out the blind extraction or blind separation using other algorithms too. The difficulty stems from the lack of solvability.

V. Stability Analysis

In this section, the results of stability analysis of model (4) and (5) is presented.

For simplicity, the functions f, g are taken as $f(s) = s^3$, $g(s) = s$, as in [4] and [6].

If $c_{12}, \cdots, c_{1n}, c_{21}, \cdots, c_{n1}$ satisfy

$$\begin{aligned} &E\left(z_1^3 z_2\right) = 0, \cdots, E\left(z_1^3 z_1\right) = 0, E\left(z_2^3 z_1\right) = 0, \cdots, \\ &E\left(z_n^3 z_1\right) = 0 \end{aligned} \tag{37}$$

then $[c_{12}, \cdots, c_{1n}, c_{21}, \cdots, c_{n1}]$ is an equilibrium of (5). From the discussion in Section III, if the extractability condition holds, then

$$c_{12} = k_2, \cdots, c_{1n} = k_n, c_{21} = \frac{a_{21}}{b}, \cdots, c_{n1} = \frac{a_{n1}}{b} \tag{38}$$

is an equilibrium of (5).

In the following, the stability of the equilibrium is analyzed. Set

$$\begin{aligned} c'_{12} &= c_{12} + \Delta c_{12}, \cdots, c'_{1n} = c_{1n} + \Delta c_{1n} \\ c'_{21} &= c_{21} + \Delta c_{21}, \cdots, c'_{n1} = c_{n1} + \Delta c_{n1}. \end{aligned} \tag{39}$$

Thus for $j = 2,, \cdots, n$

$$\frac{dc'_{1j}(t)}{dt} = \frac{d\Delta c_{1j}(t)}{dt} = \alpha_{1j} \left(z_1(t)\right)^3 z_j(t) \tag{40}$$

$$\frac{c'_{j1}(t)}{dt} = \frac{d\Delta c_{j1}(t)}{dt} = \alpha_{j1} \left(z_j(t)\right)^3 z_1(t) \tag{41}$$

and

$$\begin{cases} z_1(t) = y_1(t) - \sum_{j=2}^{n} \left(c_{1j} + \Delta c_{1j}(t)\right) y_j(t), \\ z_2(t) = y_2(t) - \left(c_{21} + \Delta, c_{21}(t)\right) \\ \qquad \cdot \left[y_1(t) - \sum_{j=2}^{n} \left(c_{1j} + \Delta c_{1j}(t)\right) y_j(t) \right], \\ \vdots \\ z_n(t) = y_n(t) - \left(c_{n1} + \Delta c_{n1}(t)\right) \\ \qquad \cdot \left[y_1(t) - \sum_{j=2}^{n} \left(c_{1j} + \Delta c_{1j}(t)\right) y_j(t) \right]. \end{cases} \tag{42}$$

Set $z_1^0(t) = y_1(t) - \sum_{j=2}^{n}(c_{1j}y_j(t))$, $z_k^0(t) = y_k(t) - c_{k1}[y_1(t) - \sum_{j=2}^{n}(c_{1j}y_j(t))]$, $k = 2, \cdots, n$, which are the outputs of the network at the equilibrium.

Expanding $(z_1(t))^3 z_k(t)$, $(z_k(t))^3 z_1(t) (k = 2, \cdots, n)$ into Taylor series and retaining the first-order term, we obtain

$$\begin{aligned} &\left(z_1(t)\right)^3 z_k(t) \\ &= \left[y_1(t) - \sum_{j=2}^{n} \left(c_{1j} + \Delta c_{1j}(t)\right) y_j(t) \right]^3 \\ &\quad \cdot \left[y_k(t) - \left(c_{k1} + \Delta c_{k1}(t)\right) \right. \\ &\quad \left. \cdot \left(y_1(t) - \sum_{j=2}^{n} \left(c_{1j} + \Delta c_{1j}(t)\right) y_j(t) \right) \right] \\ &= \left[z_1^0(t) - \sum_{j=2}^{n} \Delta c_{1j}(t) y_j(t) \right]^3 \\ &\quad \cdot \left[y_k(t) - \left(c_{k1} + \Delta c_{k1}(t)\right) \right. \\ &\quad \left. \cdot \left(z_1^0(t) - \sum_{j=2}^{n} \left(\Delta c_{1j}(t) y_j(t)\right) \right) \right] \\ &= \left(z_1^0(t)\right)^3 z_k^0(t) + c_{k1} \left(z_1^0(t)\right)^3 \\ &\quad \sum_{j=2}^{n} \cdot \left(\Delta c_{1j}(t) y_j(t) - \left(z_1^0(t)\right)^4 \right. \\ &\quad \cdot \Delta c_{k1}(t) - 3 \left(z_1^0(t)\right)^2 z_k^0(t) \\ &\quad \sum_{j=2}^{n} \cdot \left(\Delta c_{1j}(t) y_j(t)\right) + \cdots, \qquad k = 2, \cdots, n. \end{aligned} \tag{43}$$

By taking the expectations of right-hand side of (40) and (41), and noting that $E(z_1^0(t))^3 z_k^0(t) = 0$, $k = 2, \cdots, n$ from (37), the first-order approximation equations of (40) and (41) can be obtained, for $k = 2, \cdots, n$

$$\begin{aligned} &\frac{d\Delta c_{1k}(t)}{dt} \\ &\quad = \alpha_{1k} \sum_{j=2}^{n} E\left[c_{k1} \left(z_1^0(t)\right)^3 y_j(t) - 3 \left(z_1^0(t)\right)^2 z_k^0(t) y_j(t) \right] \\ &\quad \cdot \Delta c_{1j}(t) - \alpha_{1k} E\left(\left(z_1^0(t)\right)^4 \right) \Delta c_{k1}(t) \end{aligned} \tag{44}$$

$$\begin{aligned} &\frac{d\Delta c_{k1}(t)}{dt} \\ &\quad = \alpha_{k1} \sum_{j=2}^{n} E\left[3c_{k1} \left(z_k^0(t)\right)^2 z_1^0(t) y_j(t) - \left(z_k^0(t)\right)^3 y_j(t) \right] \\ &\quad \cdot \Delta c_{1j}(t) - 3\alpha_{k1} E\left(\left(z_k^0(t)\right)^2 \left(z_1^0(t)\right)^2 \right) \Delta c_{k1}(t). \end{aligned} \tag{45}$$

By combining the two equations above, a $(2n-2)$th order equations is obtained, which is denoted as

$$\begin{aligned} &\left[\frac{d\Delta c_{12}(t)}{dt}, \cdots, \frac{d\Delta c_{1n}(t)}{dt}, \frac{d\Delta c_{21}(t)}{dt}, \cdots, \frac{d\Delta c_{n1}(t)}{dt} \right]^T \\ &\quad = B\left[\Delta c_{12}(t), \cdots, \Delta c_{1n}(t), \Delta c_{21}(t), \cdots, \Delta c_{n1}(t)\right]^T \end{aligned} \tag{46}$$

where B is the corresponding coefficient matrix.

According to the well-known stability theory, if all eigenvalues of B in (46) have negative real parts, then the equilibrium composed of $c_{1j},\ c_{j1},\ j=2,\cdots,n$ is asymptotically stable.

It is not difficult to find that k_2 and k_3 in (38) are not unique, thus the equilibrium of (40) and (41) is not unique. Obviously, the stability here is a kind of local stability. Since the equilibria are not unique, if there are more than two stable equilibria, then it is impossible to ensure the global stability. In general, there exist many equilibria and some of which are not asymptotically stable. Thus it is important to choose initial values and step sizes properly for the adaptive algorithm. Yet, for many other algorithms, properly choosing initial values and step sizes is equally essential, unless the underlying cost functions are convex.

VI. Simulations Results

Simulation results presented in this section are divided into two parts. The first part of the simulation is on sequential extraction in Section IV. The second part of the simulation is concerned with the stability analysis discussed in Section V.

In order to check the performance of the sequential blind extraction in simulations, we use the following index. Denote

$$D=\begin{bmatrix}1 & -c_{12} & \cdots & -c_{1n}\\ 0 & 1 & \cdots & 0\\ \vdots & \vdots & \ddots & \vdots\\ 0 & 0 & \cdots & 1\end{bmatrix}$$

$$A=\begin{bmatrix}d_{11} & d_{12} & \cdots & d_{1n}\\ \vdots & \vdots & \ddots & \vdots\\ d_{n1} & d_{n2} & \cdots & d_{nn}\end{bmatrix}$$

where A is the mixing matrix, then the index of convergence ISI is defined as

$$\mathrm{ISI}=\frac{\sum_{i=1}^{n}|d_{1i}|-\max\{|d_{1j}|,\ j=1,\cdots,n\}}{\max\{|d_{1j}|,\ j=1,\cdots,n\}}.$$

Example 1: Consider a case to recover two sound sources $x_1(t)$ and $x_2(t)$ singularly mixed with two sources of random noises $x_3(t)$ and $x_4(t)$, where x_1 is from a speech signal of a male, x_2 is from a music signal, x_3 is an uniform white noise with values in $[-0.5,\ 0.5]$, x_4 is a Gaussian noise with mean 0.0 and variance 1.0. The underlying mixing matrix is assumed to be

$$A=\begin{bmatrix}1.5 & 0.6 & 0.3 & 0.3\\ 0 & 0 & 1 & 0.5\\ 0 & 1 & 0.5 & 0.5\\ 0 & 2 & 1 & 1\end{bmatrix}.$$

To avoid local minima and periodic oscillation and improve performance of learning as in [15], additive noise is introduced, that is

$$f(z_k(t))=[z_k(t)+10n_1(t)\exp(-0.5t)]^3,\ g(z_k(t))=z_k(t)$$

where n_1 is a Gaussian noise with zero mean and variance of 1.0.

Figs. 3 and 4 show the results of sequential blind extraction of two sources. Specifically, the first extracted signal is z_1 shown in Fig. 3, and the second extracted signal is e_1 shown in Fig. 4. The main outputs z_1 and the inputs of the second extraction $z_2,\ z_3,\ z_4$ are obtained simultaneously. From Fig. 3, we can see that the error $z_1-1.5x_1$ is very small (i.e., within $[-0.2,\ +0.2]$). The decreasing ISI is about within [0, 0.2].

In the second extraction shown in Fig. 4, the main extracted signal is z_3 instead of z_2. Fig. 4 shows that the source x_2 is recovered as e_1 up to a scale. e_2 and e_3 are unextracted noises in the second extraction.

Example 2: A speech signal of a male $x_1(t)$ is singularly mixed with two sources of random noises $x_2(t)$ and $x_3(t)$, where $x_2(t)$ is a uniform white noise with values in $[-0.5,\ +0.5]$, $x_3(t)$ is a Gaussian noise with zero mean and variance of 1.0. The mixing matrix is assumed to be

$$\begin{bmatrix}1 & 1 & 1\\ 0.5 & 0.5 & 0.5\\ 0 & 1 & 1\end{bmatrix}.$$

From the discussion in Section V, the corresponding first-order approximate equation of (40) and (41) is as follows:

$$\begin{cases}\dfrac{d\Delta c_{12}(t)}{dt}\\ \quad=\alpha_{12}\Big\{E\Big(c_{21}(z_1^0(t))^3y_2(t)-3(z_1^0(t))^2z_2^0(t)y_2(t)\Big)\\ \qquad\cdot\Delta c_{12}(t)+E\Big(c_{21}(z_1^0(t))^3y_3(t)-3(z_1^0(t))^2\\ \qquad\cdot z_2^0(t)y_3(t)\Big)\Delta c_{13}(t)-E\Big((z_1^0(t))^4\Big)\Delta c_{21}(t)\Big\}\\ \dfrac{d\Delta c_{13}(t)}{dt}\\ \quad=\alpha_{13}\Big\{E\Big(c_{31}(z_1^0(t))^3y_2(t)-3(z_1^0(t))^2z_3^0(t)y_2(t)\Big)\\ \qquad\cdot\Delta c_{12}(t)+E\Big(c_{31}(z_1^0(t))^3y_3(t)-3(z_1^0(t))^2\\ \qquad\cdot z_3^0(t)y_3(t)\Big)\Delta c_{13}(t)-E\Big((z_1^0(t))^4\Big)\Delta c_{31}(t)\Big\}\\ \dfrac{d\Delta c_{21}(t)}{dt}\\ \quad=\alpha_{21}\Big\{E\Big(\Big[3c_{21}(z_2^0(t))^2z_1^0(t)-(z_2^0(t))^3\Big]y_2(t)\Big)\\ \qquad\cdot\Delta c_{12}(t)+E\Big(3c_{21}\Big[(z_2^0(t))^2z_1^0(t)-(z_2^0(t))^3\Big]y_3(t)\Big)\\ \qquad\cdot\Delta c_{13}(t)-3E\Big((z_2^0(t))^2\Big)(z_1^0(t))^2\Big)\Delta c_{21}(t)\Big\}\\ \dfrac{d\Delta c_{31}(t)}{dt}\\ \quad=\alpha_{31}\Big\{E\Big(3c_{31}\Big[(z_3^0(t))^2z_1^0(t)-(z_3^0(t))^3\Big]y_2(t)\Big)\\ \qquad\cdot\Delta c_{12}(t)+E\Big(\Big[3c_{31}(z_3^0(t))^2z_1^0(t)-(z_3^0(t))^3\Big]y_3(t)\Big)\\ \qquad\cdot\Delta c_{13}(t)-3E\Big((z_3^0(t))^2(z_1^0(t))^2\Big)\Delta c_{31}(t)\Big\}\end{cases}. \tag{47}$$

As discussed in the preceding section, the equilibrium of the algorithm is not unique. For example,

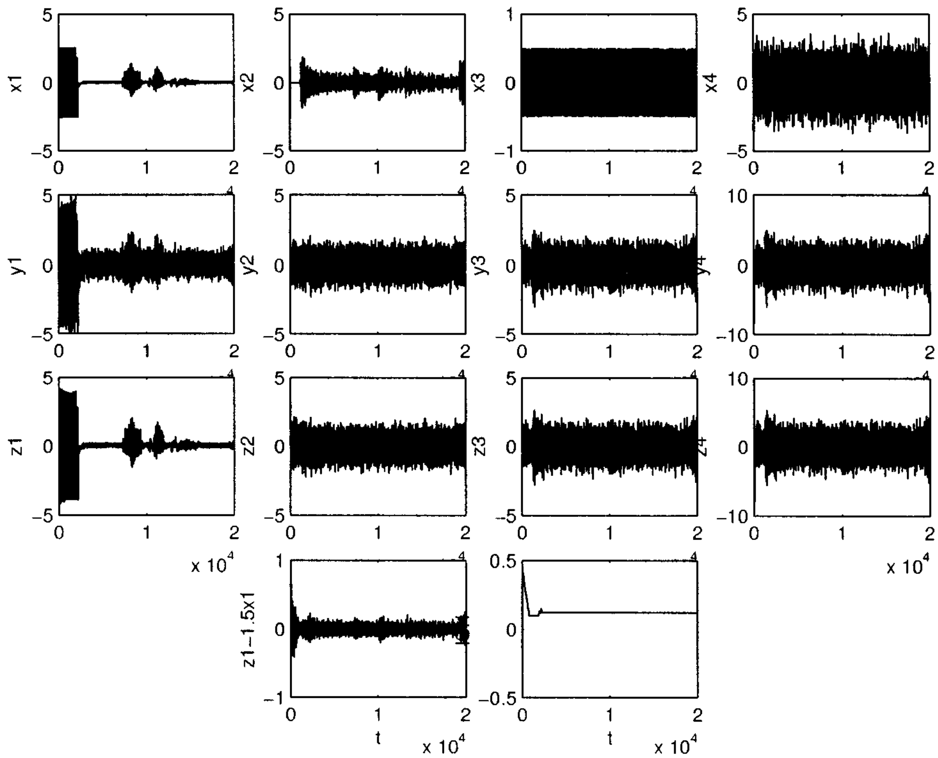


Fig. 3. The first blind extraction for ill-conditioned mixtures of four sources in Example 1.

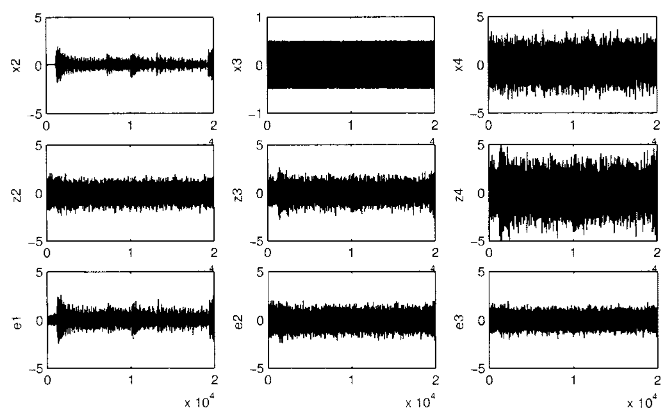


Fig. 4. The next blind extraction for ill-conditioned mixtures of remaining three sources in Example 1.

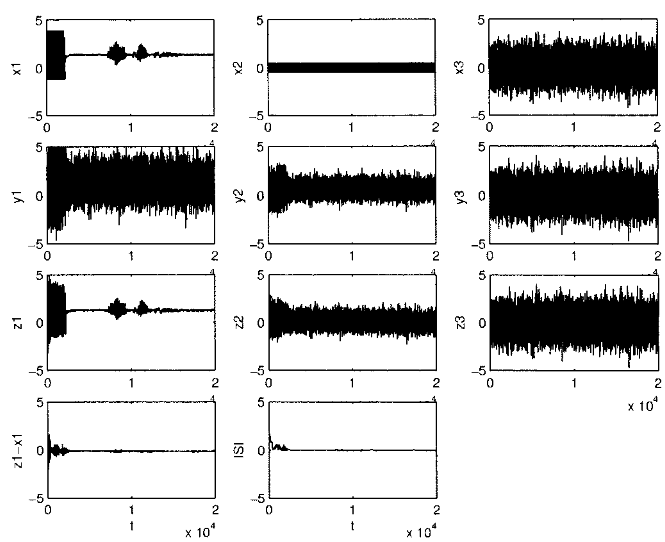


Fig. 5. Blind extraction for ill-conditioned mixtures of three sources in Example 2.

$[c_{12}, c_{13}, c_{21}, c_{31} = [0, 1, 0, 0]$ is an equilibrium, and $[c_{12}, c_{13}, c_{21}, c_{31}] = [1, 0.5, 1, 0]$ is also an equilibrium. In the simulation, the equilibrium is obtained by using the adaptive algorithm.

Set $\alpha_{12} = 0.000\,003\,9, \alpha_{13} = 0.0006, \alpha_{21} = 0.000\,05, \alpha_{31} = 0.0001$, all initial values of c_{ij} as zeros. The computed equilibrium is $[c_{12}, c_{13}, c_{21}, c_{31}] = [0.0112, 0.9741, 0.1275, 0.0747]$.

In Fig. 5, the decreasing $z_1 - x_1$ and ISI over time reveal the stability of the equilibrium above.

The proposed adaptive algorithm is also suitable for the case in which the number of sources is larger than that of the observables; i.e., A is an $n \times m$ dimensional matrix, where $m > n$, as can be seen in Example 2. Note that in Example 2, $y_1 = 2_{y2}$, in fact, they are the same signals up to a scaling factor.

VII. Concluding Remarks

We have presented analytical and simulation results on sequential blind source extractions with singular mixing matrices. We have discussed the extractability conditions, the neural-network model and its associated adaptive learning rule, and the stability results of the algorithm for ill-conditioned mixtures.

Simulation results using three and four source mixtures have been used to illustrate the proposed approach and have confirmed the validity of the theoretical results and demonstrated the performance of the algorithm.

The proposed adaptive algorithm provides a new paradigm for blind source extraction with singular mixing matrices. Since nonsingular mixing matrices always satisfy the extractability condition, the adaptive algorithm can be also used for blind sources extraction with nonsingular mixing matrices. In addition, the adaptive algorithm presented in this paper is also suitable for the case in which the number of sources is larger than that of the observables.

REFERENCES

[1] D. Kundur and D. Hatzinakos, "A novel blind deconvolution scheme for image restoration using recursive filtering," *IEEE Trans. Signal Processing*, vol. 46, pp. 375–390, 1998.

[2] D. G. Lainiotis and P. Papaparaskeva, "A partitioned adaptive approach to nonlinear channel equalization," *IEEE Trans. Commun.*, vol. 46, no. 10, pp. 1325–1336, 1998.

[3] H. H. Yang, S. I. Amari, and A. Cichocki, "Information-theoretic approach to blind separation of sources in nonlinear mixture," *Signal Processing*, vol. 64, pp. 291–300, 1998.

[4] C. Jutten and J. Herault, "Blind separation of sources, Part I: An adaptive algorithm based on neuromimetic architecture," *Signal Processing*, vol. 24, pp. 1–10, 1991.

[5] P. Common, C. Jutten, and J. Herault, "Blind separation of sources, Part II: Problems statement," *Signal Processing*, vol. 24, pp. 11–20, 1991.

[6] E. Sorouchyari, "Blind separation of sources, Part III: Stability analysis," *Signal Processing*, vol. 24, pp. 21–29, 1991.

[7] P. Comon, "Independent component analysis, a new concept?," *Signal Processing*, vol. 36, pp. 287–314, 1994.

[8] A. J. Bell and T. J. Sejnowski, "An information-maximization approach to blind separation and blind deconvolution," *Neural Comput.*, vol. 7, pp. 1004–1034, 1995.

[9] N. Delfosse and P. Loubaton, "Adaptive blind separation of independent sources: A deflation approach," *Signal Processing*, vol. 45, pp. 59–83, 1995.

[10] E. Moreau and O. Macchi, "Self-adaptive source separation-Part II: Comparison of the direct, feedback, and mixed linear network," *IEEE Trans. Signal Processing*, vol. 46, pp. 39–50, 1998.

[11] B. C. Ihm and d. J. Park, "Blind separation of sources using higher-order cumulants," *Signal Processing*, vol. 73, pp. 267–276, 1999.

[12] N. Delfosse and P. Loubaton, "Adaptive blind separation of independent sources: A deflation approach," *Signal Processing*, vol. 45, pp. 59–83, 1995.

[13] A. Hyvarinen and E. Oja, "Simple neuron models for independent component analysis," *Int. J. Neural Syst.*, vol. 7, no. 6, pp. 671–687, 1996.

[14] A. Cichocki, S. I. Amari, and R. Thawonmas, "Blind signal extraction using selfadaptive nonlinear Hebbian learning rule," in *Proc. 1996 Int. Symp. Nonlinear Theory Applicat.*, Kochi, Japan, 1996, pp. 377–380.

[15] R. Thawonmas and A. Cichocki, "Blind extraction of source signals with specified stochastic features," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Processing*, vol. 4, 1997, pp. 3353–3356.

[16] R. Thawonmas, A. Cichocki, and S. Amari, "A cascade neural network for blind extraction without spurious equilibria," *IEICE Trans. Fundamentals*, vol. E81-A, no. 9, pp. 1833–1846, 1998.

[17] A. Cichocki, "Robust neural networks with on-line learning for blind identification and blind separation," *IEEE Trans. Circuits Syst. I*, vol. 43, pp. 894–906, Nov. 1996.

[18] S. C. Douglas and A. Cichocki, "Neural networks for blind decorrelation of signals," *IEEE Trans. Signal Processing*, vol. 45, pp. 2829–2840, Nov. 1997.

[19] U. A. Lindgren and H. Broman, "Source separation using a criterion based on second-order statistics," *IEEE Trans. Signal Processing*, vol. 46, pp. 1837–1850, July 1998.

[20] L. Tong, G. H. Xu, and T. Kailath, "Blind identification and equalization based on second-order statistics: A time domain approach," *IEEE Trans. Inform. Theory*, vol. 40, pp. 340–349, 1994.

[21] Y. Inouye and K. Hirano, "Cumulant-based blind identification of linear multiinput–multioutput systems driven by colored inputs," *IEEE Trans. Signal Processing*, vol. 45, pp. 1543–1552, June 1997.

[22] H. L. N. Thi and C. Jutten, "Blind source separation for convolutive mixtures," *Signal Processing*, vol. 45, pp. 209–229, 1995.

[23] O. Shalvi and E. Weinstein, "Super-exponential methods for blind deconvolution," *IEEE Trans. Inform. Theory*, vol. 39, pp. 504–519, 1993.

[24] K. L. Yeung and S. F. Yau, "A cumulant-based super-exponential algorithm for blind deconvolution of multi-input multi-output systems," *Signal Processing*, vol. 67, pp. 141–162, 1998.

[25] M. I. Gurelli and C. L. Nikias, "EVAM: An eigenvector-based algorithm for multichannel blind deconvolution of input colored signals," *IEEE Trans. Signal Processing*, vol. 43, pp. 134–149, Jan. 1995.

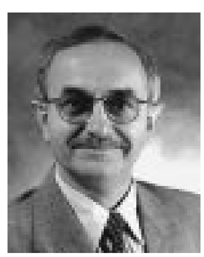
Yuanqing Li was born in Hunan Province, China, in 1966. He received the B.S. degree in applied mathematics from Wuhan University, China, in 1988, the M.S. degree in applied mathematics from South China Normal University, China, in 1994, and the Ph.D. degree in control theory and applications from South China University of Technology, China, in 1997.

Presently, he is an Associate Professor in Automatic Control Engineering Department at South China University of Technology. His research interests include applied mathematics, singular systems with delay, and blind signal processing.

Jun Wang (S'89–M'90–SM'93) received the B.S. degree in electrical engineering and the M.S. degree in systems engineering from Dalian University of Technology, China. He received the Ph.D. degree in systems engineering from Case Western Reserve University, Cleveland, OH.

He is an Associate Professor and the Director of Computational Intelligence Lab in the Department of Automation and Computer-Aided Engineering at the Chinese University of Hong Kong. Prior to this position, he was an Associate Professor at the University of North Dakota, Grand Forks, ND. His current research interests include neural networks and their engineering applications.

Dr. Wang is an Associate Editor of the IEEE TRANSACTIONS ON NEURAL NETWORKS.



Jacek M. Zurada (M'82–SM-83–F'96) is the Samuel T. Fife Alumni Professor of electrical and computer engineering at the University of Louisville, Louisville, KY. He has authored or coauthored more than 160 papers in the area of neural networks and circuits and systems for signal processing. He authored *Introduction to Artificial Neural Systems* (New York: PWS, 1992), coedited the boook *Computational Intelligence: Imitating Life* (Piscataway, NJ: IEEE Press, 1994) with R. J. Marks and C. J. Robinson, and the book *Knowledge-Based Neurocomputing* (Cambridge, MA: MIT Press, 2000) with I Cloete. He has been an invited speaker on learning algorithms, associative memory, neural networks for optimization, and hybrid techniques for rule extraction at many national and international conferences.

Dr. Zurada has been the Editor-in-Chief of the IEEE TRANSACTIONS ON NEURAL NETWORKS since 1998. He was Associate Editor of the IEEE TRANSACTIONS ON NEURAL NETWORKS from 1992 to 1997 and the IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS, PART I from 1997 to 1999 and PART II from 1993 to 1997. He was Guest Coeditor of the Special issue of the Special issue on Soft Computing of Neurocomputing in 1996 and Guest Coeditor of the Special Issue on Neural Networks and Hybrid Intelligent Models of the IEEE TRANSACTIONS ON NEURAL NETWORKS. He has received a number of awards for distinction in research and teaching. He is an IEEE Distinguished Speaker.