Available at
www.ElsevierComputerScience.com
POWERED BY SCIENCE DIRECT®

ELSEVIER

Signal Processing 84 (2004) 2245–2263

SIGNAL PROCESSING

www.elsevier.com/locate/sigpro

Blind source estimation of FIR channels for binary sources: a grouping decision approach

Yuanqing Li[a,b,*], Andrzej Cichocki[a,c], Liqing Zhang[d]

[a]Laboratory for Advanced Brain Signal Processing, RIKEN Brain Science Institute, Wako-Shi, Saitama 3510198, Japan
[b]Automation Science and Engineering Institute, Southchina University of Technology, Guangzhou, China
[c]The Department of Electrical Engineering, Warsaw University of Technology, Poland
[d]Department of Computer Science and Engineering, Shanghai Jiaotong University, Shanghai 200030, China

Received 25 May 2004; received in revised form 12 July 2004

Abstract

This paper proposes a novel grouping decision approach for blind source estimation of FIR (finite impulse response) channels with binary sources. First, solvability is discussed for single-input systems and multi-input systems. Necessary and sufficient conditions for recoverability are derived. For single-input systems, a new deterministic algorithm based on grouping and decision is proposed to recover the source up to a delay. The algorithm is easy to implement and has several advantages. For instance, when the solvability conditions are satisfied, it can be applied to cases in which: (i) the channel has zeros on the unit circle or outside of the unit circle; (ii) there are fewer sensors than sources; (iii) the source is temporarily dependent. To improve noise tolerance and reduce computational cost, the algorithm is further elaborated for highly noisy channels and high-order FIR channels, respectively. For the channels with high unimodal noise, fewer peaks appear in the probability density function (pdf) of the outputs compared to the pdf of the outputs of channels with a higher SNR. After the peaks representing cluster centers are estimated using a maximum likelihood (ML) approach, the deterministic algorithm can be used. Similar to highly noisy channels, the algorithm is also effective for high-order, exponentially decaying channels after fewer cluster centers are estimated. Furthermore, blind source estimation for multi-input systems also can be carried out as with the case of single input systems. Two deflation algorithms are presented for temporarily dependent sources and i.i.d. sources. Based on the source estimation and deflation algorithms, the sources can be obtained one by one. Finally, the validity and performance of the algorithms are illustrated by several simulation examples.

Keywords: Blind source estimation; Solvability; Single input systems; Multi-input systems; Probability density function; Clustering; Sequential blind extraction

*Corresponding author.
E-mail address: liyuan@bsp.brain.riken.jp (Y. Li).

0165-1684/$ - see front matter © 2004 Elsevier B.V. All rights reserved.
doi:10.1016/j.sigpro.2004.07.020

1. Introduction

Binary signals play important roles in pattern recognition, digital signal processing, and wireless communications, etc. When multiple binary sources are transmitted in channels, the mixtures of them are often received by sensors. In this paper we consider a dynamically mixing model described as,

$$\mathbf{x}(k) = \sum_{i=1}^{n} \sum_{p=0}^{L} \mathbf{a}_{i,p} s_i(k-p) + \mathbf{v}(k), \tag{1}$$

where $\mathbf{s}(k) = [s_1(k), \ldots, s_n(k)]^{\mathrm{T}}$ $(n \geqslant 1)$ is a vector of unknown sources with mutually independent binary components. $\forall k$, $s_j(k)$ takes one of the two known discrete values $\{d_1, d_2\}$, typically $\{0, 1\}$ or $\{-1, +1\}$ $(j = 1, \ldots, n)$. $\mathbf{x}(k) = [x_1(k), \ldots, x_m(k)]^{\mathrm{T}}$ $(m \geqslant 1)$ is an available sensor signal vector (convolutive mixture), $\mathbf{a}_{i,p} = [a_{1i,p}, \ldots, a_{mi,p}]^{\mathrm{T}}$, $i = 1, \ldots, n$, $p = 0, \ldots, L$ are unknown coefficient column vectors, $\mathbf{v} = [v_1, \ldots, v_m]^{\mathrm{T}}$ is the additive white Gaussian noise vector with mutually independent components; each component v_j has zero mean and variance σ_j^2.

The task of blind source estimation is to recover sources $s_1, \ldots, s_n$ up to an arbitrary delay, an arbitrary permutation and a scale from the observable convolutive mixture $\mathbf{x}$.

Note that if $n = 1$, then (1) is a single-input system, otherwise, it is a multi-input system. For convenience of analysis, we write the single-input system as follows:

$$\mathbf{x}(k) = \sum_{p=0}^{L} \mathbf{a}_p s(k-p) + \mathbf{v}(k), \tag{2}$$

where $\mathbf{a}_p = [a_{1,p}, \ldots, a_{m,p}]^{\mathrm{T}}$, $p = 0, \ldots, L$.

Model (1), including (2), is often encountered in data communications. One typical example is antenna array processing in which the source signals are BPSK signals [2].

Generally, the problem of estimating the sources in (1) is referred to blind deconvolution. Until now, there have been many relevant references on blind deconvolution (i.e., equalization) with finite alphabet sources. In most of these papers, an inverse filter system (equalizer) is used with its output being the recovered sources. Through optimization of different cost functions, generally based on different order statistics (e.g., SOS and HOS), many algorithms have been advanced for designing the inverse filter systems e.g., [29,30,33]. For instance, the Decision Directed algorithm [17], Sato algorithm [28], Bussgang algorithm, Godard algorithm [9], constant modulus algorithm [32], subspace algorithm [15], natural gradient algorithm [35,36,38], eigenvector algorithm [13], and almost all other existing algorithms are derived under one or more of the following conditions:

1. There are no zeros on the unit circle for the case of mixing dynamic systems; that is, the mixing systems are assumed to be nonsingular systems. If the algorithm is executed online, the mixing systems should be minimum-phase systems.
2. For MIMO systems, the sensor number is larger than the source number. The system's function matrix (frequency response) has a full column rank for almost all z [31]. If the convolutive mixing model is changed into an instantaneous mixing model, the mixing matrix is full column rank [6].
3. The sources are temporarily independent (e.g., i.i.d. sequence), or are at least temporally uncorrelated.

Recently, there have been several papers in which the first and the third conditions are relaxed. e.g., in [11], the authors discussed blind equalization of FIR channels with colored signals, assuming the sources are mutually uncorrelated and have distinct power spectra. Under the condition of an i.i.d. source, a generalized eigenvector algorithm is presented for SISO systems that could deal with the channel having several zeros on the unit circle [13]. By introducing a kind of recursive filter structure for blind deconvolution of SISO systems, the channel having several zeros close to the unit circle also can be dealt with in [14]. In [7], a geometric method has been presented for blind channel identification of SISO FIR system with binary source. This method has a drawback that cannot directly apply to MIMO systems. After the channel is identified, if the inverse filtering approach is used to estimate the source, the Condition 1 above should be satisfied.

Instantaneous blind separation of digital sources also has received attention recently.

Related works include iterative algorithm [20], ML approach [1,3,4,12] matrix factorization [2,34], deterministic approach [8], algebraic approach [16], the geometrical approach [19,24], etc. Computation complexity is a main problem for many existing deterministic methods, while global convergence is a main problem for iterative methods. Based on geometric algebra properties, many algorithms have been developed for instantaneous blind source separation [10,18,23,25–27]. These algorithms can be used for quite wide classes of sources and models, e.g., finite alphabet sources, speech sources etc., linear instantaneous mixtures and even nonlinear instantaneous mixtures, in which noise is also considered. For most of these algorithms, an important procedure is to reconstruct the mixing matrix, and then estimate the sources by use of the inverse matrix of the reconstructed matrix. Thus it is often assumed that the number of sources is equal to that of sensors.

The present paper discusses blind source estimation for single input systems and multi-input systems with binary sources. We develop a new grouping decision algorithm that addresses some of the aforementioned limitations of previous approaches. Using this approach, rather than an approach based on inverse filtering, blind source estimation can be carried out without imposing the three conditions stated above. Two sets of necessary and sufficient conditions with regard to solvability are proposed for blind source estimation for multi-input systems, and for single input systems which can be seen as a special case of multi-input systems. The solvability conditions can also be easily extended for the case of finite alphabet sources. A deterministic grouping decision algorithm is then presented for blind source estimation of a single input system in which low or no noise is present. This algorithm can also be implemented online to estimate the sources, even though the channel has zeros on the unit circle or outside the circle. However, our new approach has two drawbacks: (1) it cannot be used in cases in which high noise is present; (2) the computational burden increases exponentially as channel length increases. To overcome these drawbacks, the algorithm is extended to deal successfully with these two scenarios. Finally, a sequential blind extraction approach is proposed for multi-input dynamical systems. The extraction step can be carried out as in single input systems. Two deflation algorithms are presented for temporarily dependent sources and i.i.d. sources. Based on the source estimation and deflation algorithms, the sources can be obtained one by one.

Compared with the geometric method in [7], the approach in the present paper has better noise tolerance and less computation burden, and is suitable for MIMO systems. Since it is not necessary to identify the whole channel for estimating the sources, the strict channel identifiability conditions proposed in [7] are not necessary for our approach.

The remainder of this paper is organized as follows. The solvability analysis is presented in Section 2. Blind source estimation for single input systems follows in Section 3. Sequential blind extraction for multi-input dynamical systems is discussed in Section 4. Simulation results are presented in Section 5. The concluding remarks in Section 6 review the advantages of the proposed approach and state the remaining tasks to be studied.

2. Solvability analysis

This section analyzes solvability for blind estimation of binary sources. Two sets of necessary and sufficient conditions are derived first for single-input systems, then for multi-inputs.

We consider the following noise-free models corresponding to (1) and (2), respectively.

$$\mathbf{x}(k) = \sum_{i=1}^{n} \sum_{p=0}^{L} \mathbf{a}_{i,p} s_i(k-p), \tag{3}$$

$$\mathbf{x}(k) = \sum_{p=0}^{L} \mathbf{a}_p s(k-p). \tag{4}$$

First, we introduce two definitions.

Definition 1. 1. Model (3) is said to be well-posed, if and only if there exists a set of delays $p_1, \ldots, p_n \in \{0, \ldots, L\}$, such that, $\sum_{i=1}^{n} \sum_{p=0}^{L} \mathbf{a}_{i,p} s_i(k-p) =$

$\sum_{i=1}^{n}\sum_{p=0}^{L}\mathbf{a}_{i,p}s_i'(k-p)$ implies that $[s_1(k-p_1),\ldots,s_n(k-p_n)]^{\mathrm{T}}=[s_1'(k-p_1),\ldots,s_n'(k-p_n)]^{\mathrm{T}}$, where $\mathbf{s}$, $\mathbf{s}'$ are two binary source vectors, $k=1,2,\ldots$.

2. Model (3) is said to be partially well-posed for sources $s_{i_1},\ldots,s_{i_q}$, $i_1,\ldots,i_q\in\{1,\ldots,n\}$, if and only if there exist a set of delays $p_{i_1},\ldots,p_{i_q}\in\{0,\ldots,L\}$, such that $\sum_{i=1}^{n}\sum_{p=0}^{L}\mathbf{a}_{i,p}s_i(k-p)=\sum_{i=1}^{n}\sum_{p=0}^{L}\mathbf{a}_{i,p}s_i'(k-p)$ implies that $[s_{i_1}(k-p_{i_1}),\ldots,s_{i_q}(k-p_{i_q})]^{\mathrm{T}}=[s_{i_1}'(k-p_{i_1}),\ldots,s_{i_q}'(k-p_{i_q})]^{\mathrm{T}}$ where $k=1,2,\ldots$.

Definition 2. Model (4) is said to be well-posed, if and only if there exists a delay $p_1\in\{0,\ldots,L\}$, such that $\sum_{p=0}^{L}\mathbf{a}_p s(k-p)=\sum_{p=0}^{L}\mathbf{a}_p s'(k-p)$ implies that$s(k-p_1)=s'(k-p_1)$, where s,s' are two binary sources, $k=1,2,\ldots$.

Theorem 1. 1. *Model* (3) *is well-posed, if and only if there exist* $p_1,\ldots,p_n\in\{0,\ldots,L\}$, *such that*

$$c_{1p_1}\mathbf{a}_{1,p_1}+\cdots+c_{np_n}\mathbf{a}_{n,p_n}+\sum_{j_1\neq p_1,j_1=0}^{L}c_{1j_1}\mathbf{a}_{1,j_1}+\cdots+\sum_{j_n\neq p_n,j_n=0}^{L}c_{nj_n}\mathbf{a}_{n,j_n}\neq 0, \quad (5)$$

for all constants $c_{ij}\in\{-1,0,1\}$, $i=1,\ldots,n$, $j=0,\ldots,L$, *specifically*, $\{c_{1p_1},\ldots,c_{np_n}\}$ *has at least a nonzero entry.*

2. *Model* (4) *is well-posed, if and only if there exists a* $q\in\{0,\ldots,L\}$, *such that*

$$\mathbf{a}_q+\sum_{i\neq q,i=0}^{L}c_i\mathbf{a}_i\neq 0, \quad (6)$$

where the constants $c_i\in\{-1,0,1\}$, $i=0,\ldots,q-1,q+1,\ldots,L$.

Proof. 1. *Sufficiency*: Under the condition of (5), suppose that model (3) is not well-posed; that is, there are two sources $\mathbf{s}(k),\ldots,\mathbf{s}(k-L)$, $\mathbf{s}'(k),\ldots,\mathbf{s}'(k-L)$ with $[s_1(p_1),\ldots,s_n(p_n)]^{\mathrm{T}}\neq[s_1'(p_1),\ldots,s_n'(p_n)]^{\mathrm{T}}$, such that

$$\mathbf{x}(k)=\mathbf{x}'(k), \quad (7)$$

where $\mathbf{x}(k)$, $\mathbf{x}'(k)$ are calculated from (3) based on $\mathbf{s}$, $\mathbf{s}'$, respectively.

Then

$$\begin{aligned}\mathbf{x}(k)-\mathbf{x}'(k)=\ &\mathbf{a}_{1,0}[s_1(k)-s_1'(k)]+\cdots\\&+\mathbf{a}_{1,L}[s_1(k-L)-s_1'(k-L)]+\cdots\\&+\mathbf{a}_{n,0}[s_n(k)-s_n'(k)]+\cdots\\&+\mathbf{a}_{n,L}[s_n(k-L)-s_n'(k-L)].\end{aligned} \quad (8)$$

Set $c_{ij}=\frac{s_i(k-j)-s_i'(k-j)}{d_2-d_1}$, then $c_{ij}\in\{1,0,-1\}$, $i=1,\ldots,n$, $j=0,\ldots,L$, and $\{c_{1,p_1},\ldots,c_{n,p_n}\}$ has at least a nonzero entry.

By condition (5), we have

$$\frac{1}{d_2-d_1}(\mathbf{x}(k)-\mathbf{x}'(k))=c_{1p_1}\mathbf{a}_{1,p_1}+\cdots+c_{np_n}\mathbf{a}_{1,p_n}+\sum_{j_1\neq p_1,j_1=0}^{L}c_{1j_1}\mathbf{a}_{1,j_1}+\cdots+\sum_{j_n\neq p_n,j_n=0}^{L}c_{nj_n}\mathbf{a}_{n,j_n}\neq 0. \quad (9)$$

This is in contradiction with (7).

Thus the system is well posed. Sufficiency is proved.

Necessity: If model (3) is well-posed, then there exists a set of delays $p_1,\ldots,p_n\in\{0,\ldots,L\}$, such that $\mathbf{x}(k)=\mathbf{x}'(k)$ implies $[s_1(k-p_1),\ldots,s_n(k-p_n)]^{\mathrm{T}}=[s_1'(k-p_1),\ldots,s_n'(k-p_n)]^{\mathrm{T}}$.

Suppose that condition (5) is not satisfied; that is, there exists a set of coefficients $c_{ij}\in\{1,0,-1\}$, $i=1,\ldots,n$, $j=0,\ldots,L$, and $\{c_{1,p_1},\ldots,c_{n,p_n}\}$ has at least one nonzero entry, such that

$$c_{1p_1}\mathbf{a}_{1,p_1}+\cdots+c_{np_n}\mathbf{a}_{1,p_n}+\sum_{j_1\neq p_1,j_1=0}^{L}c_{1j_1}\mathbf{a}_{1,j_1}+\cdots+\sum_{j_n\neq p_n,j_n=0}^{L}c_{nj_n}\mathbf{a}_{n,j_n}=0.$$

Select $d_{i,j}$, $d_{i,j}'\in\{d_2,d_1\}$ such that $d_{i,j}-d_{i,j}'=c_{i,j}(d_2-d_1)$, $i=1,\ldots,n$, $j=0,\ldots,L$.

Define two sources as $s_i(k-j)=d_{i,j}, s_i'(k-j)=d_{i,j}'$, $i=1,\ldots,n$, $j=0,\ldots,L$. Since $\{c_{1,p_1},\ldots,c_{n,p_n}\}$ has at least one nonzero entry, $[s_1(k-p_1),\ldots,s_n(k-p_n)]^{\mathrm{T}}\neq[s_1'(k-p_1),\ldots,s_n'(k-p_n)]^{\mathrm{T}}$.

However,

$$\begin{aligned}\mathbf{x}(k)-\mathbf{x}'(k) &= \mathbf{a}_{1,0}[s_1(k)-s_1'(k)]+\cdots\\ &\quad+\mathbf{a}_{1,L}[s_1(k-L)-s_1'(k-L)]+\cdots\\ &\quad+\mathbf{a}_{n,0}[s_n(k)-s_n'(k)]+\cdots\\ &\quad+\mathbf{a}_{n,L}[s_n(k-L)-s_n'(k-L)]\\ &=(d_2-d_1)\sum_{i=1}^{n}\sum_{j=0}^{L}c_{ij}\mathbf{a}_{\mathbf{i}}(j)\\ &=0. \end{aligned}\tag{10}$$

Since (3) is well posed, we have $[s_1(k-p_1),\ldots,s_n(k-p_n)]^{\mathrm{T}}=[s_1'(k-p_1),\ldots,s_n'(k-p_n)]^{\mathrm{T}}$. A contradiction has occurred. This implies that condition (5) holds. Necessity is proved.

2. The proof of part 2 is similar to that of part 1 and is omitted. □

Theorem 1 can be extended to the cases in which: (1) The model (3) is partially well-posed for sources $s_{i_1},\ldots,s_{i_q}$, $i_1,\ldots,i_q\in\{1,\ldots,n\}$; (2) The source is a finite alphabet.

We have the following theorem.

Theorem 2. *System* (3) *is partially well-posed for* $\{s_{i_1},\ldots,s_{i_q}\}$, *if and only if there exists* $k_{i_1},\ldots,k_{i_q}\in\{0,\ldots,L\}$, *such that* $\sum_{i=1}^{n}\sum_{j=0}^{L}c_{ij}\mathbf{a}_{i,j}\neq 0$, *where constants* $c_{ij}\in\{-1,0,1\}$, $i=1,\ldots,n$, $j=0,\ldots,L$, *specifically*, $\{c_{i_1k_{i_1}},\ldots,c_{i_qk_{i_q}}\}$ *has at least one non-zero entry.*

The proof is similar to that of Theorem 1. □

Remarks. (1) By Theorem 1, it is possible to separate all sources, even if the number of sensors is less than the number of sources. (2) In the ill-conditioned case, in which condition (5) in Theorem 1 is not satisfied (i.e., there exist several inseparable sources in the mixtures), it is still possible to extract those separable sources according to Theorem 2.

Suppose that the sources take only finite alphabet values, say $D=\{d_1,\ldots,d_J\}$, where $d_1,\ldots,d_J$ are different real numbers. Denote the set $D_1=\{d_i-d_j|d_i,d_j\in D\}$. It can be proved that Theorem 1 is extendable for the case of finite alphabet sources, except that the coefficients c_{ij}, $c_i\in D_1$ in (5) and (6), respectively.

3. Blind deconvolution algorithm for single-input dynamical systems

Suppose that the solvability condition (6) in Theorem 1 is satisfied throughout this section. First, we consider the low-noise case when the length L of the mixing channel is not large, say $L\leqslant 10$. A deterministic grouping decision approach is proposed for blind source estimation of single input systems. The algorithm is then extended for dealing effectively with the high-noise case. Although the computational burden of the deterministic algorithm increases exponentially with respect to the channel length, its variant still works for the case in which the channel is long.

3.1. Low noise or noise free case

Since the source is binary, there are at most $2^{(L+1)}$ different output vectors of the noise free model (4), denoted as a set $\mathbf{X}_0=\{\mathbf{x}_1,\ldots,\mathbf{x}_N\}$, where $N\leqslant 2^{(L+1)}$.

Set

$$d=\min\{\|\mathbf{x}_i-\mathbf{x}_j\|_2,\ i,j=1,\ldots,N,\ i\neq j\}.\tag{11}$$

The parameter d is an important factor in analyzing noise tolerance for binary source estimation. A larger d parameter implies greater noise tolerance.

First, we present an assumption regarding low noise.

Assumption 1. The noise vector in (2) satisfies the following condition:

$$\|\mathbf{v}(k)\|_2<\frac{d}{4}.\tag{12}$$

Obviously, under condition (12), there are N different clusters with a radius less than $\frac{d}{4}$ formed by the outputs of (2). Denote these clusters as N vectors $\{\mathbf{x}_1,\ldots,\mathbf{x}_N\}$; these are the outputs of the noise-free model (4), which are the center vectors of the clusters. For convenience, rewrite these vectors as a matrix,

$$\mathbf{X} = \begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1N} \\ x_{21} & x_{22} & \cdots & x_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ x_{m1} & x_{m2} & \cdots & x_{mN} \end{bmatrix}. \tag{13}$$

Remarks 2. (1) In this subsection, two mixture vectors $\mathbf{x}(p_1)$ and $\mathbf{x}(p_2)$ are said to be in the same cluster iff

$$\|\mathbf{x}(p_1) - \mathbf{x}(p_2)\|_2 < \frac{d}{2}. \tag{14}$$

(2) One cluster center of $\{\mathbf{x}_i,\ i = 1, \ldots, N\}$ can be obtained by averaging all observable mixture vectors $\mathbf{x}(p)$ belonging to the corresponding cluster.

The following assumption is necessary for the grouping decision algorithm to be presented shortly.

Assumption 2. For Model (2), there exists a column vector $\mathbf{a}(q)$, which satisfies (6) and the following inequalities:

$$\mathbf{a}_q \neq \frac{1}{2} \sum_{p \neq q, p=0}^{L} c_p \mathbf{a}_p, \tag{15}$$

for all $c_0, \ldots, c_{q-1}, c_{q+1}, \ldots, c_L \in \{1, 0, -1\}$.

Noting that the inequality conditions in Theorems 1, 2 and Assumption 2 are similar and easy to satisfy if the channel parameters are taken randomly. Thus we think these conditions are realistic.

Now we present the grouping decision algorithm. The first step is to estimate a coefficient column vector in (2). If the column vector satisfies condition (6), then all clusters can be divided into two groups by using the vector. When we obtain an output, we can obtain the corresponding source up to a delay by identifying to which group the output belongs.

Step 1. (*Estimation*) Choose a row of the matrix $\mathbf{X}$ in (13) with at least two nonzero components assumed to be the first row, and then determine the largest and the second largest components assumed to be x_{11} and x_{12}. Set

$$\bar{\mathbf{a}}_1 = \frac{1}{d_2 - d_1} [\mathbf{x}_1 - \mathbf{x}_2].$$

It is not difficult to prove that $\bar{\mathbf{a}}_1$ is one of the coefficient columns of (2), up to a sign (see Appendix A).

Step 2. For the estimated $\bar{\mathbf{a}}_1$, we have the following proposition (see Appendix B).

Proposition 1. *If* $\bar{\mathbf{a}}_1$ *satisfies the solvability condition* (6) *and condition* (15), *then for any cluster center* $\mathbf{x}_i$, *there exists exactly one cluster center* $\mathbf{x}_j$ *that satisfies one of the following two inequalities*

$$\|\mathbf{x}_i - \mathbf{x}_j - (d_2 - d_1)\bar{\mathbf{a}}_1\|_2 < \varepsilon_0, \tag{16}$$

$$\|\mathbf{x}_j - \mathbf{x}_i - (d_2 - d_1)\bar{\mathbf{a}}_1\|_2 < \varepsilon_0, \tag{17}$$

where ε_0 *is a sufficiently small positive constant chosen in advance.*

If for any cluster center $\mathbf{x}_i$, *there exists exactly one cluster center* $\mathbf{x}_j$ *that satisfies one of* (16) *and* (17) *then go to Step* 3. *Otherwise*, *it means that* $\bar{\mathbf{a}}_1$ *does not satisfy* (6) *or* (15), *go to Step* 5;

Step 3. (*Matching and Grouping*) Match the columns of $\mathbf{X}$ in a pairwise manner. That is, if $\|\mathbf{x}_i - \mathbf{x}_j - (d_2 - d_1)\bar{\mathbf{a}}_1\| < \varepsilon_0$, then $(\mathbf{x}_i, \mathbf{x}_j)$ is defined as a pair. There exist $\frac{N}{2}$ pairs according to Criterion 1 denoted as $(\mathbf{x}_{i_1}, \mathbf{x}_{i_2}), \ldots, (\mathbf{x}_{i_{(N-1)}}, \mathbf{x}_{i_N})$.

According to the pairs above, divide $\{\mathbf{x}_i\}$ into two groups denoted as,

$$\mathbf{X}^{11} = \{\mathbf{x}_{i_1}, \mathbf{x}_{i_3}, \ldots, \mathbf{x}_{i_{(N-1)}}\},$$

$$\mathbf{X}^{12} = \{\mathbf{x}_{i_2}, \mathbf{x}_{i_4}, \ldots, \mathbf{x}_{i_N}\}.$$

Step 4. (*Deconvolution*) For an output $\mathbf{x}(k)$ of (2), determine its closest cluster center. If the cluster center belongs to $\mathbf{X}^{11}$, then set the estimation value of the source as $\bar{s}(k) = d_2$, otherwise, set $\bar{s}(k) = d_1$. Thus we obtain the source up to a delay and an exchange of d_1 and d_2 (see Appendix C);

Step 5. When $\bar{\mathbf{a}}_1$ does not satisfy Criterion 1, we match the columns of $\mathbf{X}$ in a pairwise manner, as in *Step* 3. All obtained pairs are denoted as $(\mathbf{x}_{i_1}, \mathbf{x}_{i_2}), \ldots, (\mathbf{x}_{i_{(2J-1)}}, \mathbf{x}_{i_{2J}})$, where $2J \leqslant N$. Furthermore, we can obtain two groups denoted as $\bar{\mathbf{X}}^{11}$ and $\bar{\mathbf{X}}^{12}$. Based on $\bar{\mathbf{X}}^{11}$ and the method in *Step* 1, we estimate another column vector denoted as $\bar{\mathbf{a}}_2$.

If $\bar{\mathbf{a}}_2 = \pm\bar{\mathbf{a}}_1$, continue to match the entries of $\bar{\mathbf{X}}^{11}$ using $\bar{\mathbf{a}}_2$. Repeat the process above to estimate a new column vector in (4), until a column vector different from $\pm\bar{\mathbf{a}}_1$ is obtained.

If $\bar{\mathbf{a}}_2 \neq \pm\bar{\mathbf{a}}_1$, go to *Step* 2. End.

Next, we present an illustrative example to illustrate explicitly the procedure of the above algorithm.

Example 1. Consider the following noise-free SISO model,

$$x(k) = [0.1, 0.15, 0.1][s(k), s(k-1), s(k-2)]^{\mathrm{T}}, \tag{18}$$

where the source s is valued in $\in \{-1, +1\}$.

When the number of samples is sufficiently large such that $\{s(k)\}$, $\{s(k-1)\}$ and $\{s(k-2)\}$ all cover $\{-1, +1\}$, we can obtain six cluster centers (in fact, each cluster has only one point here): $[-0.35, -0.15, -0.05, 0.05, 0.15, 0.35]$.

Using the largest and second largest centers, a coefficient can be estimated: $\bar{a}_1 = \frac{1}{2}[0.35 - 0.15] = 0.1$. And $(d_2 - d_1)\bar{a}_1 = 0.2$.

For the center 0.15, there are two centers 0.35 and -0.05 satisfies (16), thus $\bar{a}_1$ does not satisfies condition (6).

Using $\bar{a}_1$, we obtain two pairs $(0.35, 0.15)$, $(0.05, -0.15)$. Thus we obtain two groups $\{0.35, 0.05\}$, $\{0.15, -0.15\}$.

Using the first group, another coefficient can be estimated as $\bar{a}_2 = 0.15$. And $(d_2 - d_1)\bar{a}_2 = 0.3$. It is not difficult to see that $\bar{a}_2 = 0.15$ satisfies Criterion 1. Thus the solvability condition is satisfied (6).

Using $\bar{a}_2$, we can obtain three pairs $(0.35, 0.05)$, $(0.15, -0.15)$, $(-0.05, -0.35)$, and two groups $X^{11} = \{0.35, 0.15, -0.05\}$, $X^{12} = \{0.05, -0.15, -0.35\}$.

For the given mixture $x(k)$, if $x(k) \in X^{11}$, then set $\bar{s}(k) = 1$, otherwise, set $\bar{s}(k) = -1$.

In fact, it can be seen that $\bar{s}(k) = s(k-1)$.

Remarks 3. (1) Note that two zeros of system (18) are on the unit circle. Standard blind deconvolution algorithms based on a general inverse filtering approach do not work for this case. (2) From the examples in this paper, we see that only one sensor (observation) is often sufficient for blind source estimation of binary sources. If there exist more than one sensors, and one row of the mixing matrix satisfies the solvability condition given in Theorem 1, then we can choose that sensor for blind source estimation to reduce the computational burden. However, it is useful to increase the sensor number. One benefit is that the system more easily satisfies the solvability condition given in Theorem 1; another is to improve robustness with respect to noise.

There exist two limitations of the deterministic algorithm: the first is that the noise level should be low such that all cluster centers representing different outputs of noise-free model (4) are discriminated easily; the second is that the computational burden increases exponentially with respect to the tap number of channel. Thus, if the tap number is not too large, e.g., $L \leqslant 10$, the algorithm will execute quickly. In fact, as can be seen in Example 2, the algorithm is also effective for long sparse FIR channel (e.g., the number of nonzero coefficients of the FIR channel is less than 10). The tasks in the next two subsections are to extend the algorithm to the high-noise and long-channel cases.

3.2. High noise case

It is not difficult to find that if the cluster centers can be estimated correctly in advance, then the proposed deterministic algorithm still works effectively. Thus the source estimation strategy for the high-noise case is divided into two steps. The first step is to estimate the cluster centers; the second is to carry out the source estimation, as in Section 3.1. The main task of this subsection is to propose a method for estimating the cluster centers.

Under assumption of Gaussian noise, the pdf of the output can be modelled as,

$$p(x_1, \ldots, x_m; \bar{\mathbf{x}}_1, \ldots, \bar{\mathbf{x}}_N) = \sum_{i=1}^{N} p_i \prod_{j=1}^{m} \frac{1}{\sqrt{2\pi}\sigma_j} e^{-\frac{(x_j - \bar{x}_{ji})^2}{2\sigma_j^2}}, \tag{19}$$

where $\{\bar{\mathbf{x}}_i = [\bar{x}_{1i}, \ldots, \bar{x}_{mi}]^{\mathrm{T}},\ i = 1, \ldots, N\}$ are N different outputs of the high-noise model (4), p_i is the probability $P(\bar{\mathbf{x}}_i)$.

From (19), we can see that the pdf of the output of (2) has a local maximum in the cluster centers, as illustrated in the first subplot of Fig. 3 for a one-dimensional mixture.

The cluster centers $\{\bar{\mathbf{x}}_i, i = 1, \ldots, N\}$ can be estimated by solving the optimization problem

$$\max_{\bar{\mathbf{x}}_1,\ldots,\bar{\mathbf{x}}_N} p(x_1, \ldots, x_m; \bar{\mathbf{x}}_1, \ldots, \bar{\mathbf{x}}_N). \quad (20)$$

In this paper, a gradient ascent algorithm is used for solving (20). To estimate all cluster centers precisely and improve convergence performance, preprocessing of the data set of outputs is performed. That is, based on the estimated density function $p(x_1, \ldots, x_m; \bar{\mathbf{x}}_1, \ldots, \bar{\mathbf{x}}_N)$, N clusters formed by the outputs of (2) are estimated coarsely, then the geometric centers of all clusters are calculated as the initial values of the gradient ascent algorithm.

Now we present the algorithm steps for estimating the cluster centers.

Step 1. Estimate the pdf $p(x_1, \ldots, x_m; \bar{\mathbf{x}}_1, \ldots, \bar{\mathbf{x}}_N)$ of the output $\mathbf{x}$.

Suppose that there are N_0 sample points of $\mathbf{x}$ denoted as a set $\mathbf{X}$, where N_0 is sufficiently large, and that the minimum and maximum of x_i are assumed to be u_i, z_i respectively, $i = 1, \ldots, m$. Thus $\mathbf{X} \subset W = [u_1, z_1] \times \cdots \times [u_m, z_m]$. Each interval $[u_j, z_j]$ is then divided equally into M sub-intervals: $I_{ji} = [u_j + (i-1)\delta_j, u_j + i\delta_j)$, $i = 1, \ldots, M-1$, and $I_{jM} = [u_j + (M-1)\delta_j, z_j]$, where $\delta_j = \frac{z_j - u_j}{M}$, and M is a sufficiently large positive integer. Denote $w(i_1, \ldots, i_m) = I_{1i_1} \times \cdots \times I_{1i_m}$, $i_1, \ldots, i_m = 1, \ldots, M$.

By estimating the number of sample points in each interval $w(i_1, \ldots, i_m)$ denoted as $n(i_1, \ldots, i_m)$, the probability for $\mathbf{x}$ belonging to the interval can be obtained; that is, $p'(i_1, \ldots, i_m) = \frac{n(i_1, \ldots, i_m)}{N_0}$, $i_1, \ldots, i_m = 1, \ldots, M$.

The pdf is smoothed by the following filter which is often used for smoothing a function in mathematics,

$$\begin{aligned} p(i_1, \ldots, i_m) = \tfrac{1}{16}[&p'(i_1, \ldots, i_k - 2, \ldots, i_m) \\ &+ 4p'(i_1, \ldots, i_k - 1, \ldots, i_m) \\ &+ p'(i_1, \ldots, i_k, \ldots, i_m) \\ &+ 4p'(i_1, \ldots, i_k + 1, \ldots, i_m) \\ &+ p'(i_1, \ldots, i_k + 2, \ldots, i_m)], \\ &k = 1, \ldots, m. \end{aligned} \quad (21)$$

Several iterations of smoothing may be necessary sometimes. As our experience in the simulations, if the function p' is smoothed for three times, it will become very smooth.

The first subplot in Fig. 3 (see Example 4) shows the pdf of the real part of the convolutive mixture of a $4 - QAM$ source.

We also can use other algorithms, e.g. the estimation algorithms based on the kernel and the spline functions [21,22], for estimating the probability density function.

Step 2. Choose a positive constant α_0 (according to the level of noise), remove all intervals in which $p(x_1, \ldots, x_m; \bar{\mathbf{x}}_1, \ldots, \bar{\mathbf{x}}_N) < \alpha_0$ from W. Then N disjointed sub-intervals denoted as $C_1, \ldots, C_N$ remain. All samples located in these sub-intervals form N clusters. Noting that we should choose the constant α_0 such that the number (N) of sub-intervals is equal to the number of peaks of the pdf $p(x_1, \ldots, x_m; \bar{\mathbf{x}}_1, \ldots, \bar{\mathbf{x}}_N)$.

Step 3. Calculate the geometric centers of $C_1, \ldots, C_{N_1}$ denoted as $\bar{\mathbf{x}}_1(0), \ldots, \bar{\mathbf{x}}_N(0)$, which are used as the initial values for the gradient ascent algorithm. Next, start the online iteration $(i = 1, \ldots, N)$:

$$\begin{aligned} \bar{x}_{1i}(k+1) &= \bar{x}_{1i}(k) + \eta_i(k) \\ &\quad \times \frac{\partial p(x_1(k), \ldots, x_m(k); \bar{\mathbf{x}}_1, \ldots, \bar{\mathbf{x}}_N)}{\partial \bar{x}_{1i}} \\ &= \bar{x}_{1i}(k) + \frac{\eta_i(k) p_i(x_1(k) - \bar{x}_{1i}(k))}{\sigma_i^2} \\ &\quad \times \prod_{l=1}^{m} \frac{1}{\sqrt{2\pi}\sigma_l} e^{-\frac{(x_l(k) - \bar{x}_{li}(k))^2}{2\sigma_l^2}}, \\ &\quad \vdots \\ \bar{x}_{mi}(k+1) &= \bar{x}_{mi}(k) + \eta_i(k) \\ &\quad \times \frac{\partial p(x_1(k), \ldots, x_m(k); \bar{\mathbf{x}}_1, \ldots, \bar{\mathbf{x}}_N)}{\partial \bar{x}_{mi}} \\ &= \bar{x}_{mi}(k) + \frac{\eta_i(k) p_i(x_m(k) - \bar{x}_{mi}(k))}{\sigma_i^2} \\ &\quad \times \prod_{l=1}^{m} \frac{1}{\sqrt{2\pi}\sigma_l} e^{-\frac{(x_l(k) - \bar{x}_{li}(k))^2}{2\sigma_l^2}}, \end{aligned} \quad (22)$$

where $\eta_i(k)$ is a step-size of the kth iteration, $\mathbf{x}(k) = [x_1(k), \ldots, x_m(k)]^T \in C_i$; that is, we only use the outputs of (2) in C_i for the iteration of $\bar{\mathbf{x}}_i$.

The second subplot in Fig. 3 presents the iteration results in which eight centers are obtained corresponding to peaks of the first subplot.

After the cluster centers are estimated, we can use the deterministic algorithm in Section 3.1 for blind estimation of the binary source.

Remarks 4. (1) In (22), the unknown noise variance σ_j^2 can be set arbitrarily before starting the iteration. The bad estimate of σ_j^2 can be counteracted by choosing the step size parameters. We also can estimate that variance using the covariance matrix of the outputs (2), and then start the iteration. (2) If there exist N different cluster clusters, then one cluster center represents an output of the corresponding noise-free model (4), and all different outputs of the noise-free model can be estimated. (3) However, as the noise level increases, several peaks of the pdf of (19) may merge into one, and only partial cluster centers can be estimated. In this case, the algorithm is still effective, which is illustrated by Example 4.

3.3. Long-channel case

Since the number of different outputs of system (4) will increase exponentially with respect to the channel length, and since all different outputs need to be classified into two groups in the algorithm, the computational burden of the deterministic grouping decision algorithm will increase exponentially with the length of the channels.

In many practical applications, (e.g., communications), long channels are often encountered. Now we extend the algorithm for dealing with a class of long channels that have a decaying characteristic. In this subsection, a SISO system is considered. A SIMO system can be discussed similarly.

First, we have an assumption for the channel.

Assumption 3. For the FIR channel $[a_0, \ldots, a_L]$, suppose there exists a k_0, $0 \leqslant k_0 \ll L$, such that: (1) $\sum_{p=k_0+1}^{L} |a_p| < \min\{|a_0|, \ldots, |a_{k_0}|\}$; (2) $\sum_{p=k_0+1}^{L} |a_p| < \min\{|a_i - a_j|, i,j = 0, \ldots, k_0, i \neq j\}$.

For instance, the exponentially decaying channel $\{a_j = \alpha\beta^j\}$, satisfies the assumption above if $|\beta| < \frac{1}{2}$, where α, β are constants.

In fact, as in the high-noise case, we need not know exactly all the different outputs of (4). Under the conditions of Assumption 3 and sufficiently low noise, the outputs of (2) will form $2^{(k_0+1)}$ clusters apparently. It suffices to classify these clusters into two groups, and the computational burden is only related to the $2^{(k_0+1)}$ clusters. Since $k_0 \ll L$, the computational burden does not increase exponentially with the length of the channels.

In the following, we present the corresponding algorithm, of which the key point is also to estimate the centers of the clusters.

Step 1. Estimate the pdf $p(x)$ of the output x as in Section 3.2, where the definition domain of pdf is denoted as $W = [u, z]$.

The first subplot in Fig. 4 (see Example 5) shows the pdf of the real part of the convolutive mixture of a $QAM-4$ source.

Step 2. Choose a positive constant α_0 (according to the level of noise), remove all intervals in which $p(x) < \alpha_0$ from W. Then N_1 disjointed subintervals denoted as $C_1, \ldots, C_{N_1}$ remain. All samples located in these sub-intervals form N_1 clusters.

Step 3. Let the initial values be the geometric centers of $C_1, \ldots, C_{N_1}$, calculate the cluster centers using a similar gradient ascent algorithm as described in (22).

The second subplot in Fig. 4 presents the iteration results in which eight centers are obtained corresponding to peaks of the first subplot.

After the cluster centers are estimated, we can use the deterministic grouping decision algorithm in Section 3.1 for blind estimation of the binary sources. That is, these estimated centers are classified into two groups according to (16) and (17) first. Note that ε_0 in (16) and (17) is much larger than that in the short-channel case. Deconvolution is then carried out according to which group a mixture belongs.

Remark 5. In fact, even though the noise level is high and the channel taps number is large, the grouping decision algorithm can be used if a better clustering algorithm is used. Recently, many efficient algorithms for clustering have been developed (e.g., [37]).

4. Sequential blind extraction for multi-input dynamic systems

In this section, we assume that the solvability conditions imposed on multi-input systems in Theorem 1 are satisfied.

For the noise-free model (3), it is not difficult to find that there are at most $2^{n(L+1)}$ different outputs. For the noisy model (1) with low additive noise, there exist at most $2^{n(L+1)}$ different clusters. Based on these outputs of the noise-free model or the clusters of the low-noise model, the deterministic algorithm in Section 3.1 can be used directly to extract one of the sources up to a delay and an exchange of d_2 and d_1. The algorithms for the high-noise case and long-channel case in Section 3 are also suitable for single-step extraction and source estimation for the multi-input systems. Thus we omitted the theoretical analysis of single-step extraction and source estimation for multi-input systems and only present the sequential blind extraction algorithm, especially the deflation algorithm. We will propose two different deflation algorithms for temporarily dependent sources and i.i.d. sources.

4.1. Sequential extraction for temporarily dependent sources

Suppose that a channel parameter vector $\bar{\mathbf{a}}_i$ and a source $\bar{s}_i$ are obtained in the ith single-step extraction using the grouping decision algorithm in Section 3. For estimating the next channel parameter, set

$$\bar{\mathbf{x}}^{(i+1)} = \bar{\mathbf{x}}^{(i)} - \bar{\mathbf{a}}_i\bar{s}_i, \tag{23}$$

where $\bar{\mathbf{x}}^{(0)} = \mathbf{x}$.

Based on the new mixture $\bar{\mathbf{x}}^{(i+1)}$, another channel parameter $\bar{\mathbf{a}}_{i+1}$ can be estimated, as in Section 3.1. To avoid the error of the previous extraction entering into the next extraction, we always use the original mixture $\mathbf{x}$ for the $(i+1)$th extraction and obtain another source $\bar{s}_{i+1}$.

Even though it is possible that $\bar{s}_{i+1}$ is the same as one of $\bar{s}_1, \ldots, \bar{s}_i$ up to a delay, we can obtain all sources by repeating the process above at most $(n-1)L+1$.

With the sequential blind extraction presented above, all extractions are based on the original mixture, and there is no deflation process. Thus, it is unnecessary for all sources to be temporarily independent (or temporally uncorrelated).

4.2. Sequential extraction for i.i.d. sources

For model (1), if the sources $s_1, \ldots, s_n$ are an independent *i.i.d.* stochastic sequence with zero mean, the general deflation algorithm can be used in sequential extraction.

Suppose that a source $\bar{s}_1$ has been obtained using the algorithm proposed in Section 3. We use the following model to deflate the source from the mixtures $x_1, \ldots, x_m$

$$x_1^{(1)}(k) = x_1(k) - \sum_{p=-M}^{M} \bar{a}_{1,p}\bar{s}_1(k-p),$$

$$\vdots$$

$$x_m^{(1)}(k) = x_m(k) - \sum_{p=-M}^{M} \bar{a}_{m,p}\bar{s}_1(k-p), \tag{24}$$

where M is a sufficiently large positive integer which can be set by the practical problem to be dealt with $\bar{a}_{i,j},\ i = 1, \ldots, m,\ j = -M, \ldots, M$ are parameters to be determined.

Set

$$J_i = \sum_{p=-M}^{M} |\mathrm{Cum}_{2,2}(x_i^{(1)}(k), \bar{s}_1(k-p))|, \quad i = 1, \ldots, m, \tag{25}$$

where the cumulant $\mathrm{Cum}_{2,2}(x_i^{(1)}(k), \bar{s}_1(k-p)) = E\{[x_i^{(1)}(k)]^2[\bar{s}_1(k-p)]^2\} - E[x_i^{(1)}(k)]^2 E[\bar{s}_1(k-p)]^2$ in view that $s_1, \ldots, s_n$ have zero means. The deflation problem will be converted into the following m simultaneous (i.e., parallel) optimization problems

$$\min_{\bar{a}_{i,-M}, \ldots, \bar{a}_{i,M}} J_i, \quad (i = 1, \ldots, m). \tag{26}$$

Solving the optimization problems above, we can obtain the parameters $\bar{a}_{i,p}$ in Eq. (24) and carry out the deflation.

Theorem 3. (1) *If* $J_i = 0,\ i = 1, \ldots, m$, *then the new mixtures* $x_1^{(1)}, \ldots, x_m^{(1)}$ *do not contain the source* $\bar{s}_1$.

(2) *For every one of the optimization problems in* (26), *there exists a unique stable local minimum which is a global minimum.*

Proof. 1. Without loss of generality, suppose that $\bar{s}_1$ is the source s_1 up to a delay l_0; that is, $\bar{s}_1(k) = s_1(k - l_0)$. Since M is sufficiently large, the new mixtures can be represented as

$$x_1^{(1)}(k) = \sum_{j=-M+l_0}^{M+l_0} c_{1,j} s_1(k-j) + \sum_{j=2}^{n} \sum_{p=0}^{L} a_{1j}(p) s_j(k-p),$$
$$\vdots$$
$$x_m^{(1)}(k) = \sum_{j=-M+l_0}^{M+l_0} c_{m,j} s_1(k-j) + \sum_{j=2}^{n} \sum_{p=0}^{L} a_{nj}(p) s_j(k-p), \quad (27)$$

where $c_{i,j} = a_{i1}(j) - \bar{a}_{1(j-l_0)}$, $i = 1, \ldots, m$, $j = -M + l_0, \ldots, M + l_0$. Note that if $j<0$, or $j>L$, then let $a_{i1}(j) = 0$.

According to the property of cumulant and independence of $s_{i_1}(p_1)$ and $s_{i_2}(p_2)$ with $i_1 \neq i_2$ or $p_1 \neq p_2$ [5], we have

$$J_i = \sum_{p=-M+l_0}^{M+l_0} c_{i,p}^2 |\beta_1|, \quad i = 1, \ldots, m, \quad (28)$$

where $\beta_1 = \mathrm{cum}_4(s_1)$.

Obviously, $J_i = 0$, $i = 1, \ldots, m$ implies that $c_{i,p} = 0$, $i = 1, \ldots, m$, $p = -M + l_0, \ldots, M + l_0$. Thus the new mixtures $x_1^{(1)}, \ldots, x_m^{(1)}$ do not contain the source $\bar{s}_1$.

2. In view of (28), we have

$$\frac{\partial J_i}{\partial c_{i,j}} = 2c_{i,j}|\beta_1|, \quad i = 1, \ldots, m, \quad j = -M + l_0, \ldots, M + l_0.$$

Obviously, $\frac{\partial J_i}{\partial c_{i,j}} = 0$, $j = -M + l_0, \ldots, M + l_0$ have a unique solution $c_{i,j} = 0$, that is $\bar{a}_{1(j-l_0)} = a_{i1}(j)$, $j = -M + l_0, \ldots, M + l_0$.

The Hessian matrix $\frac{\partial^2 J_i}{\partial c_{i,j} \partial c_{i,k}} = \mathrm{diag}[2|\beta_1|, \ldots, 2|\beta_1|]$, which is positive definite.

Thus every optimization in (26) has a unique stable local minimum which is a global minimum. The global minimum is the true solution of deflation problem.

Now we use the gradient descent algorithm to solve the optimization problems in (26).

First calculate β_1 according to $\bar{s}_1$. If $\beta_1 > 0$, then $J_i = \sum_{p=-M}^{M} \mathrm{Cum}_{2,2}(x_i^{(1)}(k), \bar{s}_1(k-p))$, otherwise, $J_i = -\sum_{p=-M}^{M} \mathrm{Cum}_{2,2}(x_i^{(1)}(k), \bar{s}_1(k-p))$.

Without loss of generality, suppose that $\beta_1 > 0$. From (24), we have

$$J_i = \sum_{p=-M}^{M} \mathrm{Cum}_{2,2}(x_i^{(1)}(k), \bar{s}_1(k-p)) = \sum_{p=-M}^{M} [\mathrm{Cum}_{2,2}(x_i(k), \bar{s}_1(k-p)) - 2\bar{a}_{ip}\mathrm{Cum}_{1,3}(x_i(k), \bar{s}_1(k-p)) + \bar{a}_{ip}^2 \beta_1]. \quad (29)$$

Hence

$$\frac{\partial J_i}{\partial \bar{a}_{ip}} = -2\mathrm{Cum}_{1,3}(x_i(k), \bar{s}_1(k-p)) + 2\bar{a}_{ip}\beta_1, \quad (30)$$

and we obtain the following gradient algorithm

$$\Delta \bar{a}_{ip} = \eta(2\mathrm{Cum}_{1,3}(x_i(k), \bar{s}_1(k-p)) - 2\bar{a}_{ip}\beta_1), \quad p = -M, \ldots, M, \quad (31)$$

where $i = 1, \ldots, m$, η is a step size.

When a single step extraction and deflation are carried out, the result of the next extraction will be another source. This is the advantage of the deflation algorithm. However, an obvious limitation is that all sources must be temporarily independent, at least temporally uncorrelated. Many kinds of sources (e.g., image sources) do not satisfy the condition.

5. Simulation results

Simulation results presented in this section are divided into two categories. Example 2 is concerned with blind estimation of a text source that has only one convolutive mixture and additive low Gaussian noise. Example 3 considers a single-input, two-output system with a $4-QAM$ source. In Examples 4 and 5, the high-noise and long-channel cases are considered for SISO systems with a $4-QAM$ source, respectively. Example 6 concerns the sequential blind extraction for a two-input, single-output system with image sources.

Example 2. Consider the following model,

$$x(k) = [3.5, 3, 4.2, 3.5, 4.2, 7.3][s(k), s(k-4), s(k-8), s(k-12), s(k-16), s(k-22)]^{\mathrm{T}} + v(k), \qquad (32)$$

where s is a binary (black and white) text image with 250×250 pixels, v is Gaussian white noise, which satisfies Assumption 1. Of course, only the convolutive mixture x is available.

For the model (32), the length of the channel is 23. Among the 23 channel coefficients, only six coefficients are nonzero. This kind of channel is said to be sparse channel.

Fig. 1 shows the blind source estimation results in which the first subplot represents the source, the second subplot represents the convolutive output of (32), and the third subplot represents the recovered source.

Example 3. Consider the following SIMO system,

$$x_1(k) = \mathbf{a}[s(k), s(k-1), \ldots, s(k-8)]^{\mathrm{T}} + v_1,$$
$$x_2(k) = \mathbf{b}[s(k), s(k-1), \ldots, s(k-8)]^{\mathrm{T}} + v_2, \qquad (33)$$

where **a** and **b** are set randomly as [2.8904, 3.2093, 0.2677, 1.1882, 4.8813, 1.8067, 9.4199, 3.2020, 7.4138] and [3.8766, 8.0179, 3.0048, 4.5501, 4.4740, 3.5663, 5.8346, 4.8278, 7.2899], respectively (according to uniform distribution in [0,10]), $s(k)$ is a $4-QAM$ source valued randomly in $\{-1-i, -1+i, 1-i, 1+i\}$, v_1, v_2 are complex valued noises, of which all real parts and imaginary parts are $0.01n(k)$, $n(k)$ is Gaussian white noise with mean of zero and variance of 1.

System (33) can be transformed into the following two SIMO system with binary inputs,

$$\begin{cases} \mathrm{real}(x_1(k)) = & \mathbf{a}[\mathrm{real}(s(k)), \mathrm{real}(s(k-1)), \\ & \ldots, \mathrm{real}(s(k-8))]^{\mathrm{T}} + \mathrm{real}(v_1), \\ \mathrm{real}(x_2(k)) = & \mathbf{b}[\mathrm{real}(s(k)), \mathrm{real}(s(k-1)), \\ & \ldots, \mathrm{real}(s(k-8))]^{\mathrm{T}} + \mathrm{real}(v_2), \end{cases} \qquad (34)$$

$$\begin{cases} \mathrm{imag}(x_1(k)) = & \mathbf{a}[\mathrm{imag}(s(k)), \mathrm{imag}(s(k-1)), \\ & \ldots, \mathrm{imag}(s(k-8))]^{\mathrm{T}} + \mathrm{imag}(v_1), \\ \mathrm{imag}(x_2(k)) = & \mathbf{b}[\mathrm{imag}(s(k)), \mathrm{imag}(s(k-1)), \\ & \ldots, \mathrm{imag}(s(k-8))]^{\mathrm{T}} + \mathrm{imag}(v_2), \end{cases} \qquad (35)$$

where real$(\cdot)$, imag$(\cdot)$ represent the real part and imaginary part, respectively.

For (34), a coefficient column vector assumed to be $[a_j, b_j]^{\mathrm{T}}$ is estimated first, and then the source estimation is carried out using the deterministic algorithm presented in Section 3. Thus the real part of the source is obtained up to a delay, assumed to be real$(s(k-j+1))$.

For (35), the source estimation is carried out using the $[a_j, b_j]^{\mathrm{T}}$ estimated above. The imaginary part imag$(s(k-j))$ of the source is then obtained up to the same delay, as in the recovered real part.

The recovered $4-QAM$ source $\bar{s}(k) = \mathrm{real}(s(k-j+1)) + \mathrm{imag}(s(k-j+1))$.

Fig. 2 shows the blind estimation results. The $4-QAM$ source is shown in the left subplot; two outputs of (33) are shown in the second and the third subplots, respectively; the recovered source is shown in the fourth subplot. The bit error rate was calculated to be 0.001.

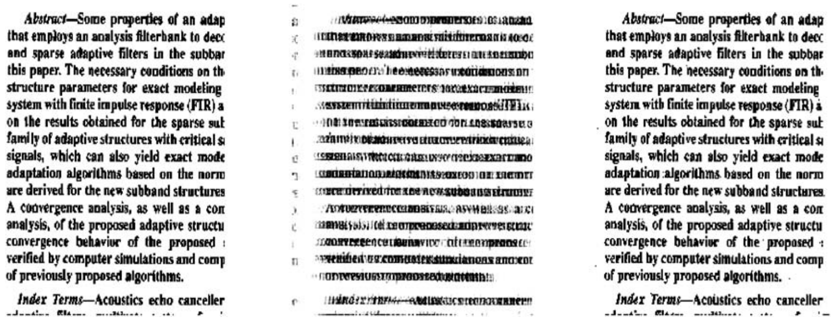

Fig. 1. Blind source estimation for SISO systems considered in Example 2. Left, black and white text image source; Middle, available convolutive mixture corrupted by low additive noise; Right, the recovered source.

The noise tolerance of the SIMO system is higher than that of its counterpart SISO system. We used the first system of (33) for blind source estimation simulation. Under the same noise situation, the bit error rate is 0.0326.

Example 4. Consider model (2) with a $4-QAM$ source and additive Gaussian, complex-valued noise. The channel parameter vector $a=[3.5, 2, 4, 3.5]$. Using the source estimation algorithm for the high-noise case considered in Section 3.2, six simulation experiments were carried out in different noise situations. Fig. 3 shows the simulation results. The left and middle subplots show the estimated pdf and iterative result of cluster centers with signal noise ratio (SNR) 18.328 dB, respectively. The right subplot shows the curve for the bit error rate with respect to SNR, calculated from the six simulations.

Note that there are 12 different outputs of the corresponding noise-free model under the channel parameter vector. With the low-noise case, there will be 12 peaks in the pdf. Because of the high noise, only eight peaks appear in the pdf, corresponding to eight clusters.

Example 5. Consider the noise free model (2). The channel parameter vector $\mathbf{a}=[16, 13, 7, 0, 0, 0.31250.9, 0, 0.0391, 0.8779, 0, 0.0049, 0.3902, 0, 0.0006, 0.1734, 0, \ldots, 0.0001]$ with length of 45, the source is a $4-QAM$ signal valued randomly in $\{-1-i, -1+i, 1-i, 1+i\}$, noise v is complex

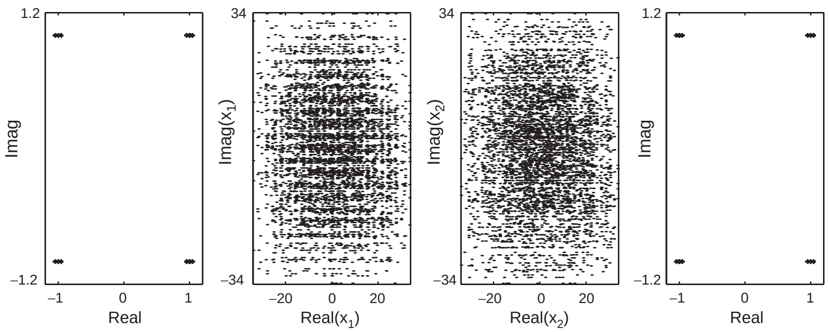

Fig. 2. Blind source estimation for two convolutive mixtures of one $4-QAM$ source considered in Example 3. Left, $4-QAM$ source; Middle two subplots, two available convolutive mixtures; Right, the recovered source.

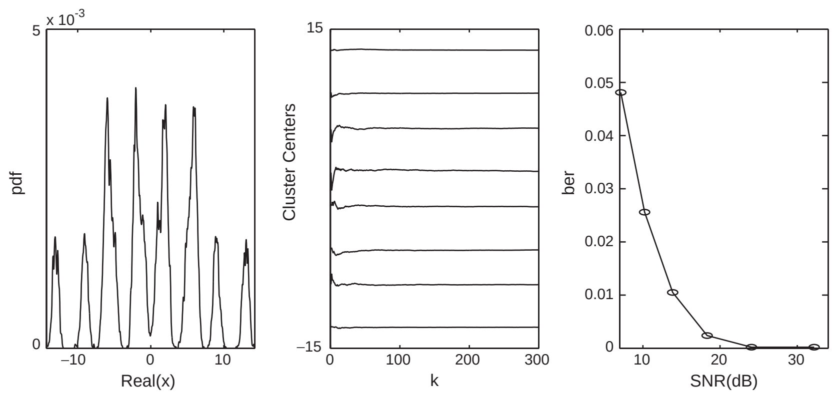

Fig. 3. Blind source estimation results in different noise situations considered in Example 4. Left, probability density function (pdf) for SNR = 18.3258 dB; Middle, iterative result of cluster centers corresponding to data in left subplot; Right, bit error rate as a function of SNR.

valued with its real and imaginary components equal to $0.01n(k)$, $n(k)$ is Gaussian white noise with mean of zero and variance of 1.

As in Example 3, the real part estimation of the source is carried out first using the real component of the mixture. In fact, the clusters of the imaginary component of the mixture are the same as those of the real component. Thus based on the imaginary component of the mixture, the imaginary part estimation of the source can be carried out according to the groups obtained in the real part estimation of the source.

Fig. 4 shows the source estimation result. In the first row, the left subplot shows the estimated pdf of the real component of the mixture, and the right subplot shows the iterative result of the cluster centers of the real component. In the second row, the left subplot shows the mixture, and the right subplot shows the recovered $4-QAM$ source. The bit error rate is 0.0001.

Example 6. Consider model (3) with two inputs of 250×250 binary text images and a single output. Channel parameters $\mathbf{a}_1$ and $\mathbf{a}_2$ are set randomly as [3.1536, 5.4027, 0.1571, 5.1255] and [5.7722, 4.0325, 2.3827, 4.0842], respectively. Using the deterministic algorithm presented in Section 3.1, a channel parameter $\bar{a}_1$ is estimated as 0.1571, and a source $\bar{s}_1$ is recovered first with a bit error rate of 0.0099.

For estimating the second channel parameter, set

$$\bar{x}(i,j) = x(i,j) - \bar{a}_1 \bar{s}_1(i,j), \tag{36}$$

where $x(i,j)$ is the mixture.

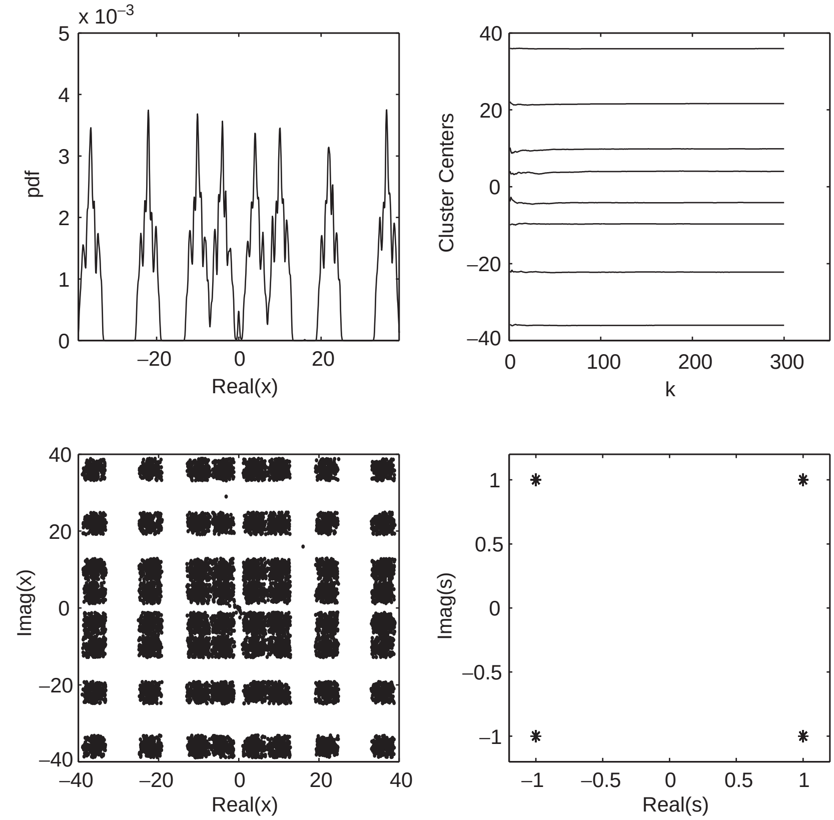

Fig. 4. Blind source estimation for one convolutive mixture of one $4-QAM$ source considered in Example 5. Top left, estimated pdf of real component of the mixture; Top right, estimated cluster centers of the real component of the mixture; Bottom left, the mixture; Bottom right, the recovered $4-QAM$ source.

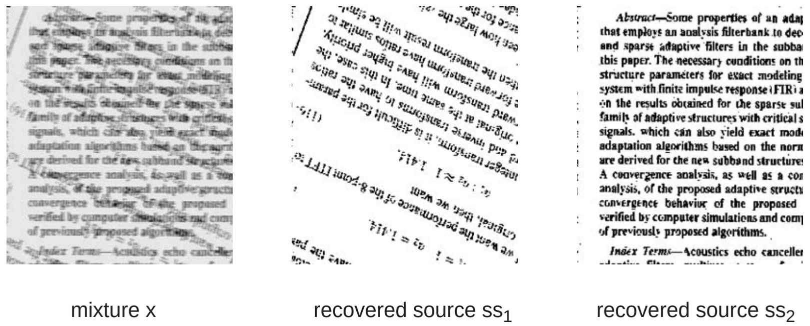

Fig. 5. Sequential blind extraction for one convolutive mixture of two text image sources considered in Example 6. Left, the mixture; Middle, source extracted after the first extraction; Right, another source obtained after the second extraction.

Based on the new mixture $\bar{x}$, another channel parameter $\bar{a}_2$ is obtained as 2.3827. Based on $\bar{a}_2$ and the original mixture x, another source $\bar{s}_2$ is obtained from the second extraction.

Fig. 5 shows the results for the sequential extraction. The left subplot shows the convolutive mixture of two text image sources; the middle and right subplots represent the sources recovered in the first and second extractions, respectively.

6. Concluding remarks

A novel approach for blind source estimation of convolutive systems with binary sources was proposed. Necessary and sufficient conditions for recoverability were established and proved. For the low-noise and noise-free cases, a deterministic grouping decision algorithm was presented for blind source estimation of single-input dynamical systems having a binary source; this approach is also suitable for multi-input systems. Compared with existing blind deconvolution algorithms generally based on inverse filtering, the grouping decision algorithm has three advantages. First, using the proposed algorithm, there is no condition imposed on the distribution of zeros of convolutive systems. Even though the system has zeros on the unit circle or outside the unit circle, the source can be recovered online. Second, the number of sensors can be less than the number of sources. Third, the algorithm can realize blind source estimation of temporarily dependent sources (e.g., image sources), even nonstationary sources.

With the algorithm, all outputs of the noise-free model (or cluster centers of the low-free model corresponding to the outputs of the noise-free model) should be obtained precisely and classified into two groups. Thus the noise tolerance is low. Although the computational burden increases exponentially with respect to the channel length, the algorithm is still very fast when the channel taps number is not too large (e.g., less than 10).

To solve the high-noise problem, we propose alternative or extended approach by estimating the pdf of the outputs. Based on the pdf, cluster segmentation is carried out, and cluster centers are obtained using the ML approach. If the noise is sufficiently low, the pdf will have the same number of peaks as the number of different outputs of the corresponding noise-free model. If the noise is high, some of the peaks will merge or disappear. Thus a cluster center in the high-noise case may represent a set of outputs rather than a single output. By classifying these cluster centers into two groups, the deterministic algorithm is still effective. Thus the tolerance to noise can be improved significantly.

For the class of situations that have a long-decaying channel, the pdf of the outputs has the similar property as that in the high-noise case:

blind source estimation can be carried out similarly. Based on an improved clustering algorithm, we think that the more complicated case of a long-decaying channel combined with high additive noise can also be dealt with effectively using the proposed approach. However, this issue is out of scope of this paper.

Although the discussion of the present paper focused mainly on single-input systems, we envisage no large obstacle for using the algorithm directly in blind source estimation of multi-input systems. Two sequential, blind extraction approaches were discussed for multi-input systems having temporarily dependent sources and i.i.d sources.

Finally, the validity and performance of the proposed algorithms were illustrated by five simulation examples.

The remaining tasks will be to extend the algorithms to the cases of finite alphabet sources and more general system models.

Acknowledgements

This initial part of this work was supported by the Excellent Young Teachers Program of MOE, P. R. China and supported by Guangdong Province Science Foundation for Research Team program under grant number 04205783. The work of third author was supported by the National Natural Science Foundation of China under grant number 60375015.

Appendix A

Without loss of generality, suppose that $a_{10} \leqslant a_{12} \leqslant \cdots \leqslant a_{1L}$, and that $a_{1i} < 0$, $a_{1(i+1)} \geqslant 0$, $d_1 < 0 < d_2$. For other cases, the proof is similar.

There are two cases, as follows.

(1) $a_{1(i+1)} > 0$. If this were the case, then the largest component of the first row of $\mathbf{X}$ in (13) is $d_1 a_{10} + \cdots + d_1 a_{1i} + d_2 a_{1(i+1)} + \cdots + d_2 a_{1L}$. Choose the corresponding column of $\mathbf{X}$ assumed to be $\mathbf{x}_1$ with the first component being the largest.

It is not difficult to find that the second largest component of the first row of $\mathbf{X}$ is one of the following components:

$$d_1 a_{10} + \cdots + d_2 a_{1i} + d_2 a_{1(i+1)} + \cdots + d_2 a_{1L},$$

$$d_1 a_{10} + \cdots + d_1 a_{1i} + d_1 a_{1(i+1)} + d_2 a_{1(i+2)} + \cdots + d_2 a_{1L}.$$

If any one of the two corresponding columns of $\mathbf{X}$ is chosen and assumed to be $\mathbf{x}_2$, then $\frac{1}{d_2 - d_1}[\mathbf{x}_1 - \mathbf{x}_2]$ is a coefficient column vector in (4) up to a sign.

(2) $a_{1(i+1)} = a_{1(i+2)} = \cdots = a_{1(i+k)} = 0$, where $1 \leqslant k \leqslant L - i$.

Without loss of generality, suppose that $k = 1$, that is, $a_{1(i+1)} = 0$, $a_{1(i+2)} > 0$, then the largest component of the first row of $\mathbf{X}$ is $d_1 a_{10} + \cdots + d_1 a_{1i} + d_2 a_{1(i+2)} + \cdots + d_2 a_{1L}$. There are two corresponding columns of $\mathbf{X}$ with the largest first component. Choose any one of them assumed to be $\mathbf{x}_1$.

The second largest component of the first row is one of the following components:

$$d_1 a_{10} + \cdots + d_2 a_{1i} + d_2 a_{1(i+2)} + \cdots + d_2 a_{1L},$$

$$d_1 a_{10} + \cdots + d_1 a_{1i} + d_1 a_{1(i+2)} + d_2 a_{1(i+2)} + \cdots + d_2 a_{1L}.$$

Also there are two corresponding columns of $\mathbf{X}$ for each of the two components above. Choose one of the two columns assumed to be $\mathbf{x}_2$ such that the second component (or other components) is closer to that of $\mathbf{x}_1$ than another. Then, $\frac{1}{d_2 - d_1}[\mathbf{x}_1 - \mathbf{x}_2]$ is a coefficient column vector of (4) up to a sign.

Appendix B

Suppose that $\bar{\mathbf{a}}_1 = \mathbf{a}_q$; we only consider the noise-free model (4). Eqs. (16) and (17) should be the following equalities

$$\|\mathbf{x}_i - \mathbf{x}_j - (d_2 - d_1)\mathbf{a}_q\|_2 = 0, \tag{37}$$

$$\|\mathbf{x}_j - \mathbf{x}_i - (d_2 - d_1)\mathbf{a}_q\|_2 = 0. \tag{38}$$

For an output $\mathbf{x}_i$, it is not difficult to prove that there exists at least one output $\mathbf{x}_j$ which satisfies one of the equalities above. Suppose that there exists another output vector assumed to be $\mathbf{x}_k$, such that the pair $(\mathbf{x}_i, \mathbf{x}_k)$ satisfies one of (37) and (38).

Without loss of generality, assume that

$$\mathbf{x}_i - \mathbf{x}_j = (d_2 - d_1)\mathbf{a}_q. \tag{39}$$

The case $\mathbf{x}_i - \mathbf{x}_j = (d_1 - d_2)\mathbf{a}_q$ can be discussed similarly.

Note that $\mathbf{x}_i - \mathbf{x}_k = (d_2 - d_1)\mathbf{a}_q$ will lead to $\mathbf{x}_k = \mathbf{x}_j$. Then we have

$$\mathbf{x}_i - \mathbf{x}_k = (d_1 - d_2)\mathbf{a}_q. \tag{40}$$

And

$$\mathbf{x}_k - \mathbf{x}_j = 2(d_2 - d_1)\mathbf{a}_q. \tag{41}$$

From model (4), $\mathbf{x}_i, \mathbf{x}_j, \mathbf{x}_k$ can be represented as follows:

$$\mathbf{x}_i = e_{10}\mathbf{a}_0 + \cdots + e_{1q}\mathbf{a}_q + \cdots + e_{1L}\mathbf{a}_L,$$
$$\mathbf{x}_j = e_{20}\mathbf{a}_0 + \cdots + e_{2q}\mathbf{a}_q + \cdots + e_{2L}\mathbf{a}_L,$$
$$\mathbf{x}_k = e_{30}\mathbf{a}_0 + \cdots + e_{3q}\mathbf{a}_q + \cdots + e_{3L}\mathbf{a}_L,$$

where $e_{p_1p_2} \in \{d_1, d_2\}$, $p_1 = 1, 2, 3$, $p_2 = 0, \ldots, L$.

There are three cases:

Case 1. $e_{2q} = e_{3q}$.

In view of (41), we have

$$\mathbf{x}_k - \mathbf{x}_j = (e_{30} - e_{20})\mathbf{a}_0 + \cdots + 0 \cdot \mathbf{a}_q + \cdots + (e_{3L} - e_{2L})\mathbf{a}_L = 2(d_2 - d_1)\mathbf{a}_q, \tag{42}$$

which is in contraction with (15) in Assumption 2.

Case 2. $e_{2q} = d_1$, $e_{3q} = d_2$.

From (41), we have

$$(e_{30} - e_{20})\mathbf{a}_0 + \cdots + (d_2 - d_1) \cdot \mathbf{a}_q + \cdots + (e_{3L} - e_{2L})\mathbf{a}_L = 2(d_2 - d_1)\mathbf{a}_q. \tag{43}$$

From (43), we can see that $\mathbf{a}_q$ does not satisfy the solvability condition (6).

Case 3. $e_{2q} = d_2$, $e_{3q} = d_1$.

In view of (39), if $\mathbf{a}_q$ satisfies (6) and (15), then we have $e_{1q} = d_2$, $e_{2q} = d_1$ according to the discussion in Appendix C. From (40), if $\mathbf{a}_q$ satisfies (6) and (15), then we have $e_{1q} = d_1$, $e_{3q} = d_2$. Thus, under condition (15) in Assumption 2, if (6) is satisfied, Case 3 cannot happen.

In view of the analysis above, under the conditions of (6) and (15), for any output $\mathbf{x}_i$, there exists only one output vector assumed to be $\mathbf{x}_j$, such that one of (37) and (38) is satisfied. Otherwise, one of the two conditions failed to be satisfied.

Hence Criterion 1 can be used to check whether $\bar{\mathbf{a}}_1$ satisfies the two conditions (6) and (15).

Appendix C

We consider only the noise-free case here. Suppose that $\bar{\mathbf{a}}_1 = \mathbf{a}_q$, which satisfies the solvability condition (6) and Assumption 2. If $(\mathbf{x}_l, \mathbf{x}_j)$ is a pair, then $\mathbf{x}_l - \mathbf{x}_j = (d_2 - d_1)\mathbf{a}_q$.

From model (4), $\mathbf{x}_l, \mathbf{x}_j$ can be represented as follows:

$$\mathbf{x}_l = e_{10}\mathbf{a}_0 + \cdots + e_{1q}\mathbf{a}_q + \cdots + e_{1L}\mathbf{a}_L,$$
$$\mathbf{x}_j = e_{20}\mathbf{a}_0 + \cdots + e_{2q}\mathbf{a}_q + \cdots + e_{2L}\mathbf{a}_L,$$

where $e_{p_1p_2} \in \{d_1, d_2\}$, $p_1 = 1, 2$, $p_2 = 0, \ldots, L$.

Thus,

$$\mathbf{x}_l - \mathbf{x}_j = (e_{10} - e_{20})\mathbf{a}_0 + \cdots + (e_{1q} - e_{2q})\mathbf{a}_q + \cdots + (e_{1L} - e_{2L})\mathbf{a}_L = (d_2 - d_1)\mathbf{a}_q. \tag{44}$$

There exist three possible cases:

1. $e_{1q} = d_2$, $e_{2q} = d_1$.
2. $e_{1q} = d_1$, $e_{2q} = d_2$.

We have

$$\mathbf{x}_l - \mathbf{x}_j = (e_{10} - e_{20})\mathbf{a}_0 + \cdots + (d_1 - d_2)\mathbf{a}_q + \cdots + (e_{1L} - e_{2L})\mathbf{a}_L = (d_2 - d_1)\mathbf{a}_q. \tag{45}$$

This implies

$$c_0\mathbf{a}_0 + \cdots + c_{q-1}\mathbf{a}_{q-1} + c_{q+1}\mathbf{a}_{q+1} + \cdots + c_L\mathbf{a}_L = 2\mathbf{a}_q, \tag{46}$$

where $c_0, \ldots, c_{q-1}, c_{q+1}, \ldots, c_L \in \{1, 0, -1\}$, which is in contradiction with (15) in Assumption 2. Thus this case will not happen.

3. $e_{1q} = e_{2q}$. From (44), we have

$$(e_{10} - e_{20})\mathbf{a}_0 + \cdots + (d_1 - d_2)\mathbf{a}_q + \cdots + (e_{1L} - e_{2L})\mathbf{a}_L = 0. \tag{47}$$

From that $d_1 - d_2 \neq 0$ and the solvability condition (6), it is impossible for this case to happen.

Therefore, the equality $\mathbf{x}_l - \mathbf{x}_j = (d_2 - d_1)\mathbf{a}_q$ implies that only case 1 will happen.

Note that if $\bar{\mathbf{a}}_1 = -\mathbf{a}_q$, then the possible case is $e_{1q} = d_1$, $e_{2q} = d_2$, and the estimation result is $e_{1q} = d_2$, $e_{2q} = d_1$. There exists an exchange of d_2 and d_1.

References

[1] K. Anand, V.U. Reddy, Maximum likelihood estimation of constellation vectors for blind separation of co-channel BPSK signals and its performance analysis, IEEE Trans. Signal Process. 45 (7) (1997) 1736–1741.

[2] K. Anand, G. Mathew, V.U. Reddy, Blind separation of multiple co-channel BPSK signals arriving at an antenna array, IEEE Signal Process. Lett. 2 (9) (1995) 176–178.

[3] A. Belouchrani, J.F. Cardoso, Maximum likelihood source separation for discrete sources, Proc. EUSIPCO 2 (September 1994) 768–771.

[4] A. Belouchrani, A. Cichocki, An adaptive algorithm for the separation of finite alphabet signals, in: Proceedings of Workshop on Signal Processing and Applications'2000, Brisbane, Australia, December, 2000.

[5] A. Cichocki, S. Amari, Adaptive Blind Signal and Image Processing, Wiley, Chichester, 2002.

[6] S. Choi, A. Cichocki, Cascade neural networks for multichannel blind deconvolution, Electron. Lett. 34 (12) (1998) 1186–1187.

[7] K.I. Diamantaras, Blind channel identification based on the geometry of the received signal constellation, IEEE Trans. Signal Process. 50 (5) (2002) 1133–1143.

[8] K.I. Diamantaras, E. Chassioti, Blind separation of N binary sources from one observation: a deterministic approach, in: Proceedings of Second International Workshop on Independent Component Analysis and Blind Signal Separation, Helsinki, Finland, June, 2000, pp. 93–98.

[9] D.N. Godard, Self-recovering equalization and carrier tracking in two-dimensional data communication systems, IEEE Trans. Commun. 28 (11) (1980) 1867–1875.

[10] O. Grellier, P. Comon, Blind separation of discrete sources, IEEE Signal Process. Lett. 5 (8) (1998) 212–214.

[11] G.B. Giannakis, Y. Hua, P. Stoica, L. Tong, Blind identification and equalization of channels driven by colored signals, in: Signal Processing Advances in Wireless and Mobile Communications, Trends in Channel Estimation and Equalization, vol. 1, Prentice Hall PTR, 2002.

[12] Bin-Chul Ihm, Dong-Jo Park, Young-Hyun Kwon, Overcomplete blind source separation of finite alphabet sources, IEICE Trans. Inf. and Syst. E84-D (1) (January 2001).

[13] B. Jelonnek, D. Boss, K.D. Kammeyer, Generalized eigenvector algorithm for blind equalization, Signal Processing 61 (1997) 237–264.

[14] A. Kumar Nandi, S.N. Anfinsen, Blind equalization with recursive filter structures, Signal Processing 80 (2000) 2151–2167.

[15] H. Liu, G. Xu, Closed-form blind symbol estimation in digital communications, IEEE Trans. Signal Process. 43 (1995) 2714–2723.

[16] Y.Q. Li, A. Cichocki, L.Q. Zhang, Blind separation and extraction of binary sources, IEICE Trans. Fund. Electron. Commun. Comput. Sci. E86-A (3) (2003) 580–590.

[17] J.E. Mazo, Analysis of decision-directed equalizer convergence, Bell Syst. Tech. J. 59 (1980) 1857–1876.

[18] A. Mansour, N. Ohnishi, C.G. Puntonet, Blind multiuser separation of instantaneous mixture algorithm based on geometrical concepts, Singal Processing 82 (8) (2002) 1155–1175.

[19] P. Pajunen, An algorithm for binary blind source separation, Helsinki University of Technology, Laboratory of Computer and Information Science Report A36, 1996.

[20] P. Pajunen, Blind separation of binary sources with less sensors than sources, in: Proceedings of 1997 International Conference on Neural Networks, Houston, Texas, USA, 1997.

[21] D.T. Pham, Blind separation of instantaneous mixtures of sources via the Gaussian mutual information criterion, Signal Processing 81 (2) (2001) 855–870.

[22] D.T. Pham, Mutual information approach to blind separation of stationary sources, IEEE Trans. Inform. Theory 48 (7) (2002) 1935–1956.

[23] C.G. Puntonet, A. Mansour, Blind separation of sources using density estimation and simulated annealing, IEICE Trans. Fund. Electron. Commun. Comput. Sci. E84-A (10) (2001) 2539–2547.

[24] C. Puntonet, A. Prieto, C. Jutten, M. Rodriguez-Alvarez, J. Ortega, A geometry-based procedure for reconstruction of n-value signals, Signal Processing 46 (1995) 267–284.

[25] C.G. Puntonet, A. Prieto, C. Jutten, M. Rodriguez, J. Ortega, Separation of sources: a geometry-based procedure for reconstruction of n-valued signals, Signal Processing 46 (3) (1995) 267–284.

[26] A. Prieto, C.G. Puntonet, B. Prieto, A neural algorithm for blind separation of sources based on geometric properties, Signal Processing 64 (3) (1998) 315–331.

[27] C.G. Puntonet, A. Mansour, C. Bauer, E. Lang, Separation of sources using simulated annealing and competitive learning, Neuro Comput. 49 (12) (2002) 39–60.

[28] Y. Sato, A method of self-recovering equalization for multi-level ampliers, IEEE Trans. Commun. 30 (1983) 162–176.

[29] O. Shalvi, E. Weinstein, New criteria for blind deconvolution of nonminimum phase systems (channels), IEEE Trans. Inform. Theory 36 (2) (1990) 312–321.

[30] O. Shalvi, E. Weinstein, Super-exponential methods for blind deconvolution, IEEE Trans. Inform. Theory 39 (1993) 504–519.

[31] J. Tugnait, Identification and deconvolution of multichannel linear non-Gaussian processes using higher order statistics and inverse filter criteria, IEEE Trans. Signal Process. 45 (3) (1997) 658–872.

[32] J.R. Treichler, M.G. Agee, A new approach to multipath correction of constant modulus signals, IEEE

Trans. Acoustics, Speech, Signal Process. 31 (1983) 349–472.
[33] J.K. Tugnait, L. Tong, Z. Ding, Single-user channel estimation and equalization, IEEE Signal Process. Magazine 17 (3) (2000) 17–28.
[34] Alle-Jan van der Veen, Analytical method for blind binary signal separation, IEEE Trans. Signal Process. 45 (4) (1997) 1078–1082.
[35] L.Q. Zhang, A. Cichocki, Blind deconvolution of dynamical systems: a state-space approach, J. Signal Processing 4 (2) (2000) 111–130.
[36] L.Q. Zhang, S. Amari, A. Cichocki, Semiparametric model and superefficiency in blind deconvolution, Signal Processing 81 (2001) 2535–2553.
[37] B. Zhang, M. Hsu, U. Dayal, K-Harmonic means—a data clustering algorithm, in: J. Roddick, K. Hornsby (Eds.), Temporal, Spatial, and Spatio-Temporal Data Mining, Springer, Berlin, 2001.
[38] L.Q. Zhang, A. Cichocki, S. Amari, Geometrical structures of FIR manifold and multichannel blind deconvolution, J. VLSI Signal Process. 31 (2002) 31–44.