PAPER *Special Section on Blind Signal Processing: ICA and BSS*

Blind Separation and Extraction of Binary Sources

Yuanqing LI[†a)], *Nonmember*, **Andrzej CICHOCKI**[††], *Regular Member*, *and* **Liqing ZHANG**[†††], *Nonmember*

SUMMARY This paper presents novel techniques for blind separation and blind extraction of instantaneously mixed binary sources, which are suitable for the case with less sensors than sources. First, a solvability analysis is presented for a general case. Necessary and sufficient conditions for recoverability of all or some part of sources are derived. A new deterministic blind separation algorithm is then proposed to estimate the mixing matrix and separate all sources efficiently in the noise-free or low noise level case. Next, using the Maximum Likelihood (ML) approach for robust estimation of centers of clusters, we have extended the algorithm for high additive noise case. Moreover, a new sequential blind extraction algorithm has been developed, which enables us not only to extract the potentially separable sources but also estimate their number. The sources can be extracted in a specific order according to their dominance (strength) in the mixtures. At last, simulation results are presented to illustrate the validity and high performance of the algorithms.

key words: *blind separation, blind extraction, solvability conditions, cluster, noise*

1. Introduction

Binary signals play important roles in pattern recognition, digital signal processing, especially wireless communications [1], [2]. When multiple binary signals are transmitted from different sources, the mixtures of them are often received by sensors. In this paper we consider a linear instantaneous mixing model described as,

$$\mathbf{x}(k) = \mathbf{A}\mathbf{s}(k) + \mathbf{v}(k), \tag{1}$$

where $\mathbf{s}(k) = [s_1(k), \cdots, s_n(k)]^T$ is an unknown n-dimensional vector of mutually independent binary source signals which can take only two discrete values $\{d_1, d_2\}$ typically $\{0, 1\}$ or $\{-1, +1\}$. $\mathbf{x}(k) = [x_1(k), \cdots, x_m(k)]^T$ is an available m-dimensional sensor (mixed) signal vector, and $\mathbf{A}$ is an unknown $m \times n$ constant mixing matrix. $\mathbf{v}(k) \in \mathbf{R}^m$ is the additive white Gaussian noise vector with zero-mean.

In (1), the sources and the mixing matrix are unknown, only the sensor signals $\mathbf{x}(k)$ are observed. The task of blind separation is to recover sources $s_1, \cdots, s_n$ from $x_1, \cdots, x_m$.

The model (1) is often encountered in data communications. One typical example is antenna array processing where the source signals are BPSK signals [2].

Previous related works include the geometrical approach [3], [4], Maximum Likelihood approach [1], [5], iterative algorithm [6]–[8], matrix factorization [2], [9], deterministic approach [11], etc. In [3], n-valued signals were considered, and the sensor number was assumed to equal the source number. In [4], by calculation of convex hull, the mixing matrix was estimated firstly, where, the sensor number was assumed to be larger than two, and any two columns of mixing matrix must be independent. The method can deal with the case of less sensors than sources, however, the algorithm is efficient only in the very low noise case, and the calculation of convex hull will become very complex for large sensor number. Maximum Likelihood method was used to estimate the mixing matrix and source signals in [1], [5], where it was assumed that a priori densities of the sources are known. The algorithm based on likelihood maximization often depends on initial values and suffers from wrong convergence. The third main approach was iterative algorithm [6]–[8], in which, the choice of starting point and algorithm convergence are key problems which limit its application as in ML approach. Matrix factorization approach (e.g., SVD) was presented in [2], [9], where the mixing matrix $\mathbf{A}$ was assumed to be full of column rank. A polynomial criterion was presented to perform blind separation of discrete sources in [10], but the mixing matrix was assumed to be column full rank, and the pdfs of all sources are known. In [11], a deterministic approach was proposed also under the condition of low noise or no noise as in [4]. Additionally, the solvability was not discussed in all references above.

In this paper, two sets of necessary and sufficient conditions on solvability are proposed firstly, which are valid for binary sources and easily extended into the

Manuscript received July 1, 2002.
Manuscript revised October 21, 2002.
Final manuscript received November 1, 2002.

[†]The author is with Laboratory for Advanced Brain Signal Processing, RIKEN Brain Science Institute, Wako-shi, 351-0198 Japan, on the leave from College of Automation Science and Engineering, South China University of Technology, China.

[††]The author is with Laboratory for Advanced Brain Signal Processing, RIKEN Brain Science Institute, Wako-shi, 351-0198 Japan, on the leave from Warsaw University of Technology, Poland.

[†††]The author is with the Department of Computer Science and Engineering, Shanghai Jiaotong University, Shanghai 200030, China.

a) E-mail: liyuan@bsp.brain.riken.go.jp

case of finite alphabet sources. One of the two sets of conditions is for blind separation, another is for blind extraction for ill-conditioned case in which there exist several inseparable sources in the mixtures. Next, a deterministic blind separation algorithm is presented in the case of low noise or no noise. Although the algorithm is similar to that in [11], the difference between the two algorithms is in the two facts:
(1) The estimation of mixing matrix and labelling clusters are carried out simultaneously in this paper.
(2) The basic algorithm presented in this paper has also resulted in an extended blind extraction algorithm in which the extraction order of the sources can be determined in advance.

Furthermore, our extraction algorithm can deal with ill-conditioned mixtures. By exploiting probability density function, the proposed blind separation algorithm enables to estimate binary sources even if noise is relatively high. In this paper, no other conditions on binary sources but their mutual independence are assumed.

In contrast to standard blind signal separation with analogue signals where occur two kinds of ambiguities: arbitrary scaling and arbitrary permutation, in the binary blind separation, the scaling problem is reduced to binary sign ambiguity. Furthermore, we can extract sources according to their strength in the mixture.

This paper is organized as follows. The solvability analysis is presented in Sect. 2. Blind separation and extraction algorithms follow in Sects. 3 and 4 respectively. Simulation results are presented in Sect. 5. The concluding remarks in Sect. 6 review the advantages of the proposed approach and states the main limitations.

2. Solvability Analysis

This section analyzes solvability for blind separation and blind extraction of binary sources. Several sets of necessary and sufficient conditions are derived.

Rewrite mixing matrix $\mathbf{A} = [\mathbf{a}_1, \mathbf{a}_2, \cdots, \mathbf{a}_n]$, where $\mathbf{a}_i$'s are nonzero column vectors of $\mathbf{A}$, then the corresponding noise free model of (1) can be represented as follows:

$$\mathbf{x}(k) = \mathbf{a}_1 s_1(k) + \cdots + \mathbf{a}_n s_n(k). \tag{2}$$

At first, we introduce two definitions.

Definition 1: 1. The model (2) is said to be well-posed, if and only if $\mathbf{As} = \mathbf{As}'$ implies that $\mathbf{s} = \mathbf{s}'$, where $\mathbf{s}, \mathbf{s}'$ are two binary source vectors.
2. The model (2) is said to be partially well-posed for sources $s_{i_1}, \cdots, s_{i_L}$, if and only if $\mathbf{As} = \mathbf{As}'$ implies that $s_{i_j} = s'_{i_j}, j = 1, \cdots, L$, where $\mathbf{s}, \mathbf{s}'$ are two binary source vectors.

Definition 2: Given constant column vectors $\mathbf{a}_1, \mathbf{a}_2, \cdots, \mathbf{a}_n$, if there exist at least one nonzero numbers $c_1, \cdots, c_n \in \{1, 0, -1\}$, such that

$$c_1\mathbf{a}_1 + \cdots + c_n\mathbf{a}_n = 0, \tag{3}$$

then it is said that these vectors are bi-dependent (in binary sense). Otherwise, they are bi-independent.

Theorem 1: The model (2) is well-posed, if and only if all column vectors of $\mathbf{A}$ are bi-independent, i.e., $c_1\mathbf{a}_1 + \cdots + c_n\mathbf{a}_n \neq 0$ for $c_1, \cdots, c_n \in \{1, 0, -1\}$ and in such case all binary sources can be recovered.

Proof:
Sufficiency: Suppose that there are two source vectors $[s_1(k), \cdots, s_n(k)]^T$, $[s'_1(k), \cdots, s'_n(k)]^T$, such that

$$\mathbf{a}_1 s_1(k) + \cdots + \mathbf{a}_n s_n(k) = \mathbf{a}_1 s'_1(k) + \cdots + \mathbf{a}_n s'_n(k). \tag{4}$$

Then

$$\mathbf{a}_1[s_1(k) - s'_1(k)] + \cdots + \mathbf{a}_n[s_n(k) - s'_n(k)] = 0. \tag{5}$$

Set $c_i(k) = \frac{s_i(k) - s'_i(k)}{d_2 - d_1}$, then $c_i \in \{1, 0, -1\}, i = 1, \cdots, n$. By the condition of the theorem, we have

$$c_i(k) = 0,\ i = 1, \cdots, n,\ k = 1, 2, \cdots, \tag{6}$$

that is, $\mathbf{s} = \mathbf{s}'$. Sufficiency is proved.
Necessity: Suppose that the columns $\mathbf{a}_1, \cdots, \mathbf{a}_n$ are bi-dependent. Without loss of generality, assume that

$$\mathbf{a}_1 + \mathbf{a}_2 = 0. \tag{7}$$

For given source $\mathbf{s}(k) = [s_1(k), \cdots, s_n(k)]^T$, there must be some time k', such that $s_1(k') = d_1, s_2(k') = d_1$. Set

$$s'_1(k') = d_2,\ s'_2(k') = d_2,\ s'_i(k') = s_i(k'),\quad i = 3, \cdots, n. \tag{8}$$

Thus we obtain two different signals $\mathbf{s}, \mathbf{s}'$ which satisfy

$$\mathbf{a}_1[s_1(k') - s'_1(k')] + \cdots + \mathbf{a}_n[s_n(k') - s'_n(k')] = (d_1 - d_2)\mathbf{a}_1 + (d_1 - d_2)\mathbf{a}_2 = 0. \tag{9}$$

That is, $\mathbf{A}s = \mathbf{A}s'$, which is in contradiction with that (1) is well-posed. The necessity is proved.

Theorem 1 can be extended to the following two cases: 1. Only part of columns of $\mathbf{A}$ satisfy the condition in Theorem 2. Finite alphabet source case.

We have the following theorem.

Theorem 2: System (2) is partially well-posed for $\{s_{i_1}, \cdots, s_{i_L}\}$, if and only if the column vectors $\{\mathbf{a}_{i_1}, \cdots, \mathbf{a}_{i_L}\}$ are bi-independent with each other and with others of $\{\mathbf{a}_i, i = 1, \cdots, n\}$.

The proof is similar to that of Theorem 1.

Remarks 1: 1. By Theorem 1, it is possible to separate all sources even if the sensor number is less than source number. 2. In the ill-conditioned case where the condition 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 as can be seen in Sect. 4.2 and Example 4.

Suppose that the sources are taken finite values in $D = \{d_1, \cdots, d_J\}$, where $d_1, \cdots, d_J$ are different real numbers. Denote the set $D_1 = \{d_i - d_j | d_i, d_j \in D, i \neq j\}$. We have the following theorem on finite alphabet sources which is generalization of Theorem 1:

Theorem 3: System (2) is well-posed, if and only if column vectors of **A** satisfy the following inequality:

$$c_1\mathbf{a}_{i_1} + \cdots + c_L\mathbf{a}_{i_L} \neq 0, \tag{10}$$

for any $c_k \in D_1, k = 1, \cdots, L$, and $\{i_1, \cdots i_L\}$ covers all subsets of $\{1, 2, \cdots, n\}$.

The proof is similar to that of Theorem 1 and is omitted here.

3. Blind Separation Algorithm

In this section, the low noise (high signal-noise ratio (SNR)) case including noise free case is considered firstly, and a deterministic blind separation algorithm is proposed. For high noise case, the deterministic algorithm is also efficient but then the cluster centers must be estimated correctly by other methods (e.g., pdf estimation in this paper) in advance.

Suppose that the condition of Theorem 1 is always satisfied in this section. Since sources are binary, there are 2^n different output vectors of the noise free model (2), denoted as $\{\bar{\mathbf{x}}_1, \cdots, \bar{\mathbf{x}}_N\}$, where $N = 2^n$. Set

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

In this paper, the norm $||\cdot||$ refers to $||\cdot||_2$.

The parameter d is an important factor in analyzing noise tolerance for binary source separation. In this paper, the following parameter is used as a criterium for noise tolerance instead of standard SNR,

$$NDR = \frac{\sigma_v^2}{d^2}, \tag{12}$$

where σ_v^2 is the sum of variance of all noise components.

The following two subsections discuss low noise case and high noise case respectively.

3.1 Low Noise or Noise Free Case

At first, we present an assumption on low noise.

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

$$||\mathbf{v}(k)|| < \frac{d}{4}. \tag{13}$$

Obviously, under the condition (13), there are $N = 2^n$ different clusters with radius less than $\frac{d}{4}$ formed by the outputs of (1) represented by N vectors $\{\mathbf{x}_1, \cdots, \mathbf{x}_N\}$, which are the center vectors of the clusters. For convenience, rewrite these vectors as a matrix,

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

Remarks 2: 1. In this section, two mixture vectors $\mathbf{x}(k_1)$ and $\mathbf{x}(k_2)$ are said to be in the same cluster iff

$$||\mathbf{x}(k_1) - \mathbf{x}(k_2)|| < \frac{d}{2}. \tag{14}$$

2. The center of a cluster can be obtained by averaging all vectors of the cluster.

The assumption below supposed to be satisfied in this paper is necessary for the following algorithm.

Assumption 2: For the model (1), the columns of the mixing matrix **A** satisfy the following inequalities:

$$\mathbf{a}_i \neq \frac{1}{2}[c_1\mathbf{a}_1 + \cdots + c_{i-1}\mathbf{a}_{i-1} + c_{i+1}\mathbf{a}_{i+1} + c_n\mathbf{a}_n],$$
$$i = 1, \cdots, n, \tag{15}$$

where $c_1, \cdots, c_{i-1}, c_{i+1} \cdots, c_n \in \{1, 0, -1\}$.

In fact, (15) is satisfied with probability one.

Now we present the following blind separation algorithm by which the columns of mixing matrix **A** are estimated, centers of the clusters are labelled as a column vector $[\bar{s}_1, \cdots \bar{s}_n]$, where $\bar{s}_i \in \{d_1, d_2\}$, and all sources are recovered.

(The first grouping)

1. Choose a row of the matrix **X** with at least two nonzero components assumed to be the first row, and then find out the biggest and the second biggest 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 columns of **A** up to a sign (see Appendix A).

2. Given $\bar{\mathbf{a}}_1$, we match the columns of **X** pairwisely. That is, if $||\mathbf{x}_i - \mathbf{x}_j - (d_2 - d_1)\bar{\mathbf{a}}_1|| < \epsilon_0$, where ϵ_0 is a sufficiently small constant chosen in advance, then $(\mathbf{x}_i, \mathbf{x}_j)$ is defined as a pair. It can be proved that under Assumption 2, for any $\mathbf{x}_i$, there exists only one matched cluster center (see Appendix B).

There exist $\frac{N}{2}$ pairs denoted as $(\mathbf{x}_{i_1}, \mathbf{x}_{i_2})$, $\cdots, (\mathbf{x}_{i_{(N-1)}}, \mathbf{x}_{i_N})$.

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

$$\mathbf{X}^{11} = [\mathbf{x}_{i_1}, \mathbf{x}_{i_3}, \cdots, \mathbf{x}_{i_{(N-1)}}],$$
$$\mathbf{X}^{12} = [\mathbf{x}_{i_2}, \mathbf{x}_{i_4}, \cdots, \mathbf{x}_{i_N}].$$

4. For all columns of $\mathbf{X}^{11}$, set the first components of their labelling vectors to be d_2, and for all columns of $\mathbf{X}^{12}$, set the first components of their labelling vectors to be d_1 (the reason can be seen in Appendix B).

(The second grouping)

5. Repeating the above processes 1, 2 and 3 by replacing **X** with $\mathbf{X}^{11}$, we obtain another column of **A**

up to a sign denoted as $\bar{\mathbf{a}}_2$, and two groups denoted as $\mathbf{X}^{21}$ and $\mathbf{X}^{22}$. For all columns of $\mathbf{X}^{21}$, set the second components of their labelling vectors to d_2, and set the second components of the labelling vectors to d_1 for any cluster centers of $\mathbf{X}^{22}$;

6. Using $\bar{\mathbf{a}}_2$, we divide all columns of $\mathbf{X}^{12}$ into two groups denoted as $\mathbf{X}^{23}$ and $\mathbf{X}^{24}$. For all columns of $\mathbf{X}^{23}$, set the second components of their labelling vectors to d_2, and set the second components of the labelling vectors to d_1 for any cluster centers of $\mathbf{X}^{24}$;

Repeating the above process, we can divide each of $\mathbf{X}^{21}$, $\mathbf{X}^{22}$, $\mathbf{X}^{23}$ and $\mathbf{X}^{24}$ into two groups, obtain another column of $\mathbf{A}$ up to a sign and the third components of all labelling vectors. After the $n-th$ grouping, we can obtain all columns of $\mathbf{A}$ denoted as $[\bar{\mathbf{a}}_1, \cdots, \bar{\mathbf{a}}_n]$ up to a sign and order. Also we can obtain the labelling vectors of all cluster centers.

Separation)
Given a mixture vector $\mathbf{x}(k)$, find out the closest cluster center, then the corresponding labelling vector $\bar{\mathbf{s}}(k)$ is the source vector $\mathbf{s}(k)$ up to a exchange of d_1 and d_2, and source order. $k = 1, 2, \cdots$.

Here we give an illustrative example to show the procedure of the above algorithm.

Example 1: Consider the following noise free model with one mixture and four sources,

$$x(k) = [0.1, 0.15, 0.3, 1][s_1(k), s_2(k), s_3(k), s_4(k)]^T,$$

where $s_1, \cdots, s_4$ are four independent sources valued in $\{-1, +1\}$, and the coefficient vector is assumed to be unknown.

When the number of samples is sufficiently large, we can obtain 16 cluster centers (in fact, each cluster has only one point here):

$$[-1.55, -1.35, -1.25, -1.05, -0.95, -0.75, -0.65, -0.45, 0.45, 0.65, 0.75, 0.95, 1.05, 1.25, 1.35, 1.55].$$

The first grouping:

$$\bar{a}_1 = \frac{1}{2}[1.55 - 1.35] = 0.1.$$

By using $\bar{a}_1$, we can obtain all pairs:

$$\{-1.35, -1.55\}, \{-1.05, -1.25\}, \{-0.75, -0.95\}, \{-0.45, -0.65\}, \{1.55, 1.35\}, \{1.25, 1.05\}, \{0.95, 0.75\}, \{0.65, 0.45\}.$$

Thus two groups of cluster centers and the first component of labelling vectors can be obtained as follows:

$$\mathbf{X}^{11} = [-1.35, -1.05, -0.75, -0.45, 0.65, 0.95, 1.25, 1.55], \quad \bar{s}_1 = 1.$$

$$\mathbf{X}^{12} = [-1.55, -1.25, -0.95, -0.65, 0.45, 0.75, 1.05, 1.35], \quad \bar{s}_1 = -1.$$

The second grouping:

From $\mathbf{X}^{11}$, we can calculate out another component of the mixing vector, by which we obtain four center groups, and the second component of labelling vectors:

$\bar{a}_2 = \frac{1}{2}[1.55 - 1.25] = 0.15$. And the four groups are:

$$\begin{aligned}
\mathbf{X}^{21} &= [-1.05, -0.45, 0.95, 1.55], \ \bar{s}_1 = 1, \bar{s}_2 = 1, \\
\mathbf{X}^{22} &= [-1.35, -0.75, 0.65, 1.25], \ \bar{s}_1 = 1, \bar{s}_2 = -1 \\
\mathbf{X}^{23} &= [-1.25, -0.65, 0.75, 1.35], \ \bar{s}_1 = -1, \bar{s}_2 = 1, \\
\mathbf{X}^{24} &= [-1.55, -0.95, 0.45, 1.05], \ \bar{s}_1 = -1, \bar{s}_2 = -1.
\end{aligned}$$

The third grouping:

Repeating the process above, we obtain another component of the mixing vector, eight groups, and the third components of labelling vectors of all cluster centers:

$\bar{a}_3 = \frac{1}{2}[1.55 - 0.95] = 0.3$.

$$\begin{aligned}
&[-0.45, 1.55], \ \bar{s}_1 = 1, \bar{s}_2 = 1, \bar{s}_3 = 1, \\
&[-1.05, 0.95], \ \bar{s}_1 = 1, \bar{s}_2 = 1, \bar{s}_3 = -1, \\
&[-0.75, 1.25], \ \bar{s}_1 = 1, \bar{s}_2 = -1, \bar{s}_3 = 1, \\
&[-1.35, 0.65], \ \bar{s}_1 = 1, \bar{s}_2 = -1, \bar{s}_3 = -1, \\
&[-0.65, 1.35], \ \bar{s}_1 = -1, \bar{s}_2 = 1, \bar{s}_3 = 1, \\
&[-1.25, 0.75], \ \bar{s}_1 = -1, \bar{s}_2 = 1, \bar{s}_3 = -1, \\
&[-0.95, 1.05], \ \bar{s}_1 = -1, \bar{s}_2 = -1, \bar{s}_3 = 1, \\
&[-1.55, 0.45], \ \bar{s}_1 = -1, \bar{s}_2 = -1, \bar{s}_3 = -1.
\end{aligned}$$

By the last grouping, we obtain the mixing vector and all labelling vectors of cluster centers:

$$\begin{aligned}
&\mathbf{A} = [0.1, 0.15, 0.3, 1], \\
&1.55: \ [1, 1, 1, 1]^T; \quad -0.45: \ [1, 1, 1, -1]^T; \ \cdots; \\
&-1.55: \ [-1, -1, -1, -1]; \ 0.45: \ [-1, -1, -1, 1]^T.
\end{aligned}$$

After finding the correspondence between the cluster center and the labelling vectors, we recover sources in the following way. For a mixture $\mathbf{x}(k)$, find out the matched cluster center. Then the corresponding labelling vector is the source vector up to a sign and order.

Obviously, the minimal distance d of any two cluster centers in this example is 0.05. If there exists noise, then the absolute value of noise should be less than 0.0125, otherwise the algorithm here may fail to estimate all sources.

Remarks 3: 1. As illustrated in the example, only one sensor is often sufficient for blind separation 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 separation in order to reduce computation burden. The choice is based on the two facts: the first is that the outputs of the selected sensor can form $N(= 2^n)$ different clusters, which implies that the corresponding mixing coefficients satisfy the solvability condition. The second is that if there exist k sensors of which each produces N different clusters with centers $\{x_{i_j 1}, \cdots, x_{i_j N}\}, j = 1, \cdots, k$, and set

$$d_j = \min\{|x_{i_j l_1} - x_{i_j l_2}|, 1 \le l_1, l_2 \le N, l_1 \ne l_2\} \qquad (16)$$

$(j = 1, \cdots, k)$, and $d_J = \max_{1 \le j \le k}\{d_j\}$, then the J-th sensor can be chosen to recover the sources since it has the largest noise tolerance.

2. However, it is useful to increase the sensor number sometimes. One benefit is to easily satisfy the solvability condition given in Theorem 1, another is to improve noise tolerance provided that the increase of minimal distance of cluster centers is larger than that of the norm of noise vectors brought by increased mixture space dimension.

3.2 High Noise Case

It is not difficult to find that if the cluster centers are estimated correctly in advance, then the proposed deterministic algorithm can be successfully used even for noisy data. Thus the separation strategy for high noise case is to divide it into two steps: The first is to estimate cluster centers, the second one is to separate sources as in Sect. 3.1. The main task of this subsection is to propose a robust method of estimating the cluster centers.

Without loss of generality, consider the model (1) with one dimensional output (the coefficient row vector denoted as $\mathbf{a}$). The probability density function of the sensor signal can be proved to be,

$$p(x) = \sum_{i=1}^{N} \frac{p_i}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\bar{x}_i)^2}{2\sigma^2}}, \qquad (17)$$

where $\{\bar{x}_i, i = 1, \cdots, N\}$ are N different values of $\mathbf{as}$, and p_i is the probability $P(\mathbf{as} = \bar{x}_i)$. The parameter NDR should be bounded such that the pdf $p(x)$ has N different peaks. If NDR is very large, several peaks of probability density function will disappear.

From (17), we can see that the probability density function of the sensor signal of (1) has local maximum in the centers of the clusters as illustrated in the first subplot of Fig. 1 on one dimensional mixture.

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

$$\max_{\bar{x}_1, \cdots, \bar{x}_N} p(x). \qquad (18)$$

In this paper, gradient ascent algorithm is used for solving (18). To estimate all cluster centers precisely and improve convergence performance, a preprocessing of the data set of outputs is used here. That is, based on the estimated density function $p(x)$, N clusters formed by the outputs of (1) are obtained coarsely, then the geometric centers of all clusters are calculated as the initial values of the gradient ascent algorithm.

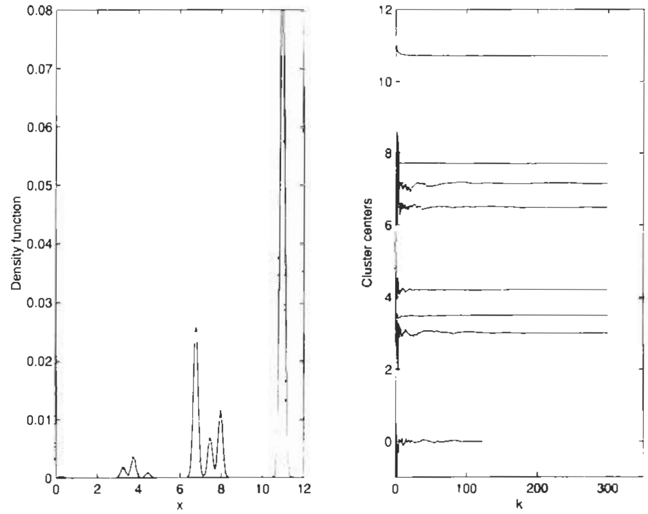

Fig. 1 The first subplot: the probability density function of outputs, the second subplot, the estimated cluster centers, where. NDR = 12.9 dB, M = 300, the sources and mixture are from Example 3.

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

Step 1. Estimate the pdf of the output x.

Suppose that there are N_0 sample points of x denoted as a set X, where $N_0 >> N$, and that the minimum and maximum in X are assumed to be a, b respectively. The interval $[a, b]$ is then divided equally into M subintervals which are $[a + i\delta, a + (i+1)\delta)$, $i = 0, \cdots, M-2$, and $[a + (M-1)\delta, b]$, where $\delta = \frac{b-a}{M}$, and M is a sufficiently large positive integer. By estimating the number of sample points in each interval denoted by n_i for the i-th interval, the probability for x belonging to the i-th interval can be obtained, that is, $p'_i = \frac{n_i}{N_0}$, $i = 1, \cdots, M$.

The pdf is smoothed by the following filter (to repeat the smoothing for several times may be more useful sometimes),

$$p_k = \frac{1}{16}[p'_{k-2} + 4p'_{k-1} + 6p'_k + 4p'_{k+1} + p'_{k+2}]. \qquad (19)$$

The first subplot in Fig. 1 shows the probability density function with one mixture of three text image sources and Gaussian noise (NDR = 12.9 dB, M = 300), where the three sources can be seen in Example 2.

Step 2. Divide coarsely the interval $[a, b]$ into N disjointed sub-intervals denoted as $C_1, \cdots, C_N$ such that every sub-interval contains only one peak of the estimated density function $p(x)$.

Step 3. Calculate the geometric centers of $C_1, \cdots, C_N$ denoted as $\bar{x}_1(0), \cdots, \bar{x}_N(0)$, which are used as the initial values of gradient ascent algorithm. Then start the iteration $(i = 1, \cdots, N)$:

$$\bar{x}_i(k+1) = \bar{x}_i(k) + \alpha_i(k)\frac{\partial p(x(k))}{\partial \bar{x}_i}$$

$$= \bar{x}_i(k) + \frac{\alpha_i(k)(x(k) - \bar{x}_i(k))}{\sqrt{2\pi}\sigma^3} \cdot e^{-\frac{(x(k)-\bar{x}_i(k))^2}{2\sigma^2}}, \quad (20)$$

where $\alpha_i(k)$ is a step-size of the k-th iteration, $x(k) \in C_i$, that is, we only use the outputs of (1) in C_i for the iteration of $\bar{x}_i$.

The second subplot in Fig. 1 presents the iteration results in which 8 centers are obtained corresponding to peaks of the first subplot.

After the cluster centers are estimated, we can use the deterministic algorithm in Sect. 3.1 for blind separation of the binary sources.

Remarks 4: 1. If the cluster centers are estimated successfully by using the set X containing sufficiently rich output sample points, the labelling of the centers are completed by the deterministic algorithm, then blind separation of sources can be carried out on line. 2. In (20), the noise variance σ_v^2 can be set arbitrarily in advance.

As illustrated in Examples 2 and 3, the algorithm in this section can be used for binary image enhancement. Furthermore, for one image with N grey levels corrupted by noise, the mixture is one dimensional and the N gray levels are the cluster centers. If the N grey levels are known, the separation step of the algorithm in Sect. 3.1 can be used directly for de-noising. Otherwise, the algorithm in Sect. 3.2 can be used for estimating the N gray levels and then noise is reduced by separation. The algorithm in this section also can be extended to deal with only one RGB color image corrupted by noise, where the mixture is three dimensional.

4. Blind Extraction Algorithm

In the separation approach, all sources are separated simultaneously. In many applications, we are interested not in separation of all sources but rather in extraction of certain sources containing useful information. Until now, there have existed many references for blind extraction of analogue sources (e.g., [12]), in which, the source signals are extracted one by one.

This section presents the blind extraction algorithm of binary sources. Two cases are considered: 1. completely well-posed case in which the solvability condition in Theorem 1 is satisfied; 2. ill-conditioned cases in which the solvability condition only in Theorem 2 not in Theorem 1 is satisfied.

In this section, the algorithm is deterministic which is suitable for low noise mixture. As in Sect. 3, after the cluster centers are obtained by estimating the pdf of outputs, the following deterministic extraction algorithm can be used in high noise case.

4.1 Completely Well-Posed Case

The blind extraction algorithm in this subsection is also based on grouping.

1. Using the similar approach as in Sect. 3.1, find out several columns of the mixing matrix $\mathbf{A}$.

2. The first extraction.

Choose a column $\bar{\mathbf{a}}_{i_1}$ firstly, divide the cluster set $\mathbf{X}$ into two sets denoted as $\mathbf{X}^{11}, \mathbf{X}^{12}$ by the similar manner as in blind separation algorithm.

For given mixture vector $\mathbf{x}(k)$, set $s_{i_1}(k) = d_2$ if $\mathbf{x}(k) \in \mathbf{X}^{11}$, otherwise, set $s_{i_1}(k) = d_1$. Thus one source is extracted.

3. The second extraction.

Set $\mathbf{x}'(k) = \mathbf{x}(k) - \bar{\mathbf{a}}_{i_1} s_{i_1}(k) + \bar{\mathbf{a}}_{i_1} d_1$, and choose a column $\bar{\mathbf{a}}_{i_2} \neq \bar{\mathbf{a}}_{i_1}$. By using $\bar{\mathbf{a}}_{i_2}$, the cluster set $\mathbf{X}_{11}$ is divided into two groups denoted as $\mathbf{X}^{21}, \mathbf{X}^{22}$ as in step 2. For new mixture vector $\mathbf{x}'(k)$, set $s_{i_2}(k) = d_2$ if $\mathbf{x}'(k) \in \mathbf{X}_{21}$, otherwise, set $s_{i_2}(k) = d_1$. Then the second source is extracted.

Using the new mixture $\mathbf{x}'(k)$, and the left columns of $\mathbf{A}$, we can continue the extraction until all interesting sources are extracted.

Remarks 5: 1. Obviously, we can determine which source to be extracted and the extraction order in advance according to the estimated columns of the mixing matrix; 2. A mixture vector belongs to a cluster set iff it satisfies the similar condition to (14).

4.2 Ill-Conditioned Case

If only the solvability condition in Theorem 2 (not in Theorem 1) is satisfied, then there are less than $N = 2^n$ clusters formed by the outputs of Model (1). For instance, if s_1, s_2 are binary sources valued in $\{-1, 1\}$, then $x = s_1 + s_2$ is valued in $\{-2, 0, 2\}$. Suppose that Model (1) produces N_1 different clusters with the centers $\{\mathbf{x}_1, \cdots, \mathbf{x}_{N_1}\}$.

1. By using the centers $\{\mathbf{x}_1, \cdots, \mathbf{x}_{N_1}\}$, a column of mixing matrix $\mathbf{A}$ is estimated firstly denoted by $\bar{\mathbf{a}}_1$ as in the blind separation algorithm in Sect. 3.1.

2. As in Sect. 3.1, we match the center vectors above pairwisely using $\bar{\mathbf{a}}_1$. There exist two cases:

a) If all centers are matched, then they can be divided into two groups and a source can be extracted by the similar manners to that in Sect. 4.1. And the next extraction can be continued.

b) If there exist several centers left when the matching is finished, then the source corresponding to $\bar{\mathbf{a}}_1$ can not be extracted. We can obtain two groups from these matched pairs although several centers are not contained in the two groups. Based on one of the two groups, another column of $\mathbf{A}$ can be estimated denoted as $\bar{\mathbf{a}}_2$. Using $\bar{\mathbf{a}}_2$, if we can re-match the centers $\{\mathbf{x}_1, \cdots, \mathbf{x}_{N_1}\}$ without any centers left, then a extraction can be carried out. Otherwise, we should continue to find out the other columns of $\mathbf{A}$ until a source is extracted.

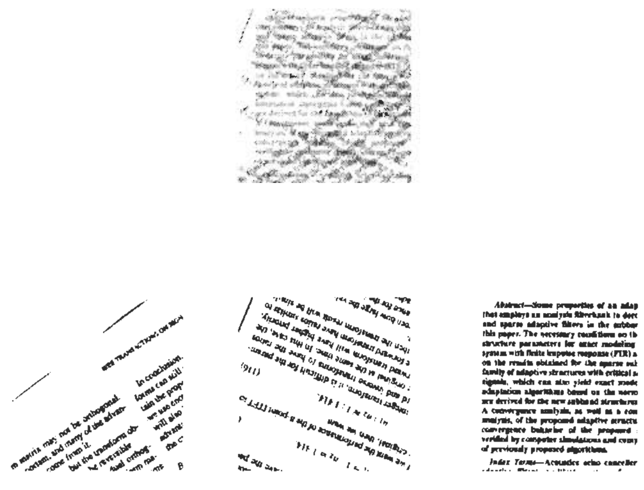

Fig. 2 Blind separation for one mixture of three text sources in Example 2. The first row: available image corrupted by other text images, the second row: three reconstructed texts.

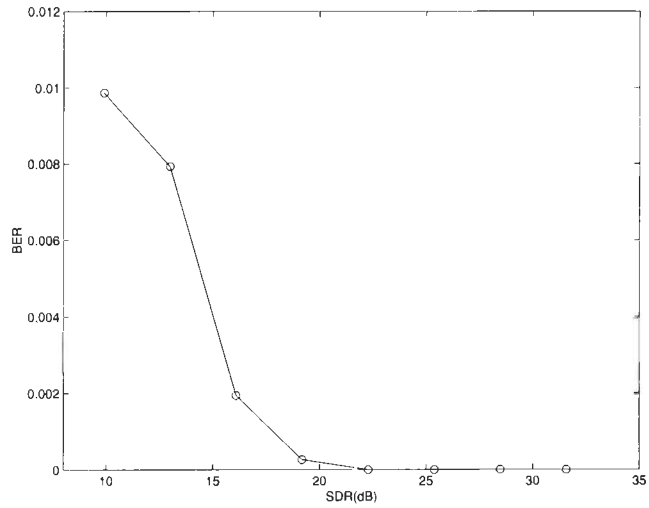

Fig. 3 The curve of bit error rate vs. NDR in Example 3.

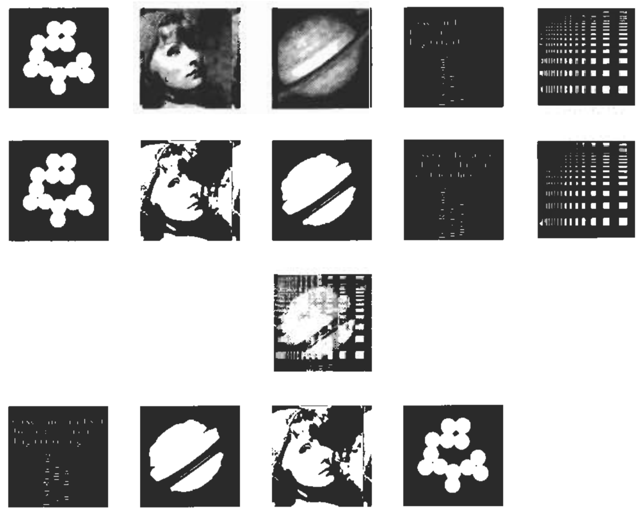

Fig. 4 The sequential blind extraction of four sources with one sensor and five image sources in Example 4. The first row: original images of which the second and the third are gray images, the others are black and white. The second row: five binary images dithered from the first row. The third row: the single mixture, the fourth row: the sequentially extracted binary images according to the descending order of mixing coefficients.

5. Simulation Results

Simulation results presented in this section are divided into four categories. Example 2 is concerned with blind separation of three text sources with only one mixture and additive low Gaussian noise. Example 3 considers the model in Example 2 with high noise, in which a curve of bit error rate vs NDR is obtained. In Example 4, well-posed case and ill-conditioned case are discussed respectively. In the well-posed case, sequential blind extraction results of 5 black-white binary image sources are presented with one mixture. The extraction is based on the descending order of the estimated mixing coefficients. In the ill-conditioned case, only one separable source is extracted. In Example 5, an application of the algorithm is presented for blind enhancement of digital images.

Example 2: Consider the following synthetic model,

$$x(i,j) = [3.5, 3.0, 4.2][s_1(i,j), s_2(i,j), s_3(i,j)]^T + v(i,j), \quad (21)$$

where s_1, s_2, s_3 are three binary text sources with 250×250 pixels, $\mathbf{v}$ is Gaussian white noise which satisfies Assumption 1, is about 25 dB on $d^2 = 0.25$. Of course, it is assumed that only the mixture x is available in the blind separation.

Figure 2 shows the blind separation results, in which, the first subplot is the mixture, the three subplots in the second row are recovered source text images. Since the recovered sources have no obvious difference with the original ones, the original sources are not presented here.

Example 3: Consider the model in Example 2 with high noise. At first, the cluster centers are estimated by using the algorithm in Sect. 3, separation of sources is then carried out by the deterministic algorithm.

By using eight different zero mean Gaussian noises with different variance. that is, with different NDRs. eight simulations are run. In each experiment. the bit error rate is calculated. Figure 3 shows the curve of bit error rate vs NDR.

Example 4: Consider noise-free mixing model with one output of five binary image sources $[s_1, \cdots, s_5]^T$ with 250×250 pixels. The mixing vector is set by $[0.1, 0.3, 0.6, 1.4, 3.2]$. According to the blind extraction algorithm in Sect. 3, the mixing vector is estimated firstly. Then sequential blind extraction can be carried out as selected extraction order in advance.

Figure 4 shows the extraction results. The first row of Fig. 4 presents five source images of which two

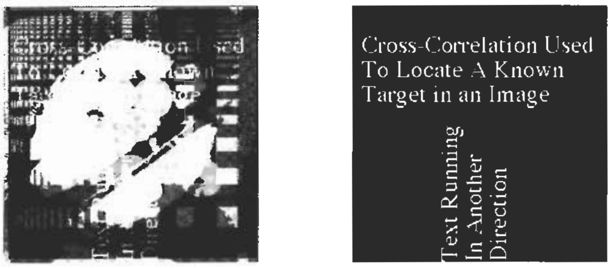

Fig. 5 The blind extraction of single source from single mixture of five binary images in the case when only one separable source is successfully extracted. The first: mixed image, the second: extracted text image.

are gray scale ones, the others are black and white. The second row shows five black white images dithered from the first row. The third row is the single mixture. Four extracted black white images are displayed in the fourth row, and the extraction order is based on the descending order of the estimated coefficients $a(1), \cdots, a(4)$.

Now set the mixing row vector $\mathbf{a}$ to be $[0.4, 0.4, 0.5, 0.7, 0.5]$. Obviously, the condition of Theorem 2 is satisfied, and there exists only one source which can be extracted. By trying several times, we can obtain the correct component of $\mathbf{a}$ for blind extraction and carry out the blind extraction. Figure 5 shows the blind extraction result, where the first subplot is the mixture, the second is the extracted source.

6. Concluding Remarks

A novel approach for blind separation and extraction of linearly mixed binary sources is proposed in this paper. Necessary and sufficient conditions on recoverability of binary sources are obtained. The result is also extended for the mixtures of finite alphabet sources. A deterministic blind binary source separation algorithm is proposed for low noise or no noise case, in which the estimating mixing matrix and labelling clusters are carried out simultaneously. For high noise case, the blind separation algorithm also can be used successfully after cluster centers are estimated based on Maximum Likelihood Approach. Sequential blind extraction approach is also discussed in this paper, in which the number and order of sources for extraction can be determined in advance according to the estimated mixing coefficients. The simulation results confirm the validity of the approach in this paper.

There are several points to be emphasized as the concluding remarks of this paper.

1) The noise tolerance of blind separation and extraction algorithm depends on the minimal distance of 2^n different output vectors of the corresponding noise free model (2). Although only one mixture is used in the simulations. increasing number of sensors can improve the noise tolerance provided that the increase of the minimal distance is larger than that of noise norm, and may easier satisfy the solvability condition given in Theorem 1.

2) If the number of sources n is not large, e.g., $n \leq 10$, the cluster number is also not large, the algorithms in this paper are very efficient, by which the real-time separation and extraction are possible.

If the source n is very large, it is difficult to obtain all different clusters, the computation time of the algorithms in this paper will increase exponentially. The open problem is how to reduce computational complexity.

3) When only part of sources are separable, that is, only the condition in Theorem 2 is satisfied, blind extraction algorithm is able to extract all sources which are separable. It reveals an advantage of blind extraction over blind separation.

References

[1] K. Anand and V.U. Reddy, "Maximum likelihood estimation of constellation vectors for blind separation of co-Channel BPSK signals and its performane analysis," IEEE Trans. Signal Process., vol.45, no.7, pp.1736–1741, 1997.

[2] K. Anand, G. Mathew, and V.U. Reddy, "Blind separation of multiple co-channel BPSK signals arriving at an antenna array," IEEE Signal Process. Lett., vol.2, no.9, pp.176–178, 1995.

[3] C. Puntonet, A. Prieto, C. Jutten, M. Rodriguez-Alvarez, and J. Ortega, "A geometry-based procedure for reconstruction of n-value signals," Signal Processing, vol.46, pp.267–284, 1995.

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

[5] A. Belouchrani and J.F. Cardoso, "Maximum likelihood source separation for discrete sources," Proc. EUSIPCO, vol.2, pp.768–771, 1994.9.

[6] P. Pajunen, "Blind separation of binary sources with less sensors than sources," Proc. 1997 International Conference on Neural Networks, Houston, Texas, USA, 1997.

[7] B.C. Ihm, D.J. Park, and Y.H. Kwon, "Overcomplete blind source separation of finite alphabet sources," IEICE Trans. Inf. & Syst., vol.E84-D, no.1, pp.209–212, Jan. 2001.

[8] A. Belouchrani and A. Cichocki, "An adaptive algorithm for the separation of finite alphabet signals," Proc. Workshop on Signal Processing and Applications 2000, Brisbane Australia, Dec. 2000.

[9] A.-J. van der Veen, "Analytical method for blind binary signal separation," IEEE Trans. Signal Process., vol.45, no.4, pp.1078–1082, 1997.

[10] O. Grellier and P. Comon, "Blind separation of discrete sources," IEEE Signal Process. Lett., vol.5, no.8, pp.212–214, 1998.

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

[12] Z. Malouche and O. Macchi, "Adaptive unsupervised extraction of one component of a linear mixture with a single neuron," IEEE Trans. Neural Netw., vol.9, no.1, pp.123–135, 1998.

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

[14] K. Matsuoka, M. Ohya, and M. Kawamoto, "A neural net for blind separation of nonstationary signals," Neural Netw., vol.8, no.3, pp.411–419, 1995.
[15] N. Takagi, "A modular inversion hardware algorithm with a redundant binary representation," IEICE Trans. Inf. & Syst., vol.E76-D, no.8, pp.863–869, Aug. 1993.

Appendix A

Without loss of generality, suppose that $a_{11} \leq a_{12} \leq \cdots \leq a_{1n}$, and that $a_{1i} < 0, a_{1(i+1)} \geq 0$, $d_1 < 0 < d_2$. For other cases, the proof is similar.

There are two cases as follows.

1) $a_{1(i+1)} > 0$. Then the biggest component of the first row of $\mathbf{X}$ is $d_1 a_{11} + \cdots + d_1 a_{1i} + d_2 a_{1(i+1)} + \cdots + d_2 a_{1n}$. Choose the corresponding column of x assumed to be $\mathbf{x}_1$ with the first component being the biggest.

It is not difficult to find that the second biggest component of the first row is one of the following components:

$$d_1 a_{11} + d_1 a_{12} + \cdots + d_2 a_{1i} + d_2 a_{1(i+1)} + \cdots + d_2 a_{1n},$$
$$d_1 a_{11} + \cdots + d_1 a_{1i} + d_1 a_{1(i+1)} + d_2 a_{1(i+2)} + \cdots + d_2 a_{1n}.$$

If any one of the two corresponding columns of $\mathbf{X}$ is chosen assumed to be $\mathbf{x}_2$, then $\frac{1}{d_2 - d_1}[\mathbf{x}_1 - \mathbf{x}_2]$ is a column of mixing matrix $\mathbf{A}$ up to a sign.

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

Without loss of generality, suppose that $k = 1$, that is, $a_{1(i+1)} = 0, a_{1(i+2)} > 0$, then the biggest component of the first row of $\mathbf{X}$ is $d_1 a_{11} + \cdots + d_1 a_{1i} + d_2 a_{1(i+2)} + \cdots + d_2 a_{1n}$. There are two corresponding columns of $\mathbf{X}$ with the biggest first component. Choose any one of them assumed to be $\mathbf{x}_1$.

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

$$d_1 a_{11} + d_1 a_{12} + \cdots + d_2 a_{1i} + d_2 a_{1(i+2)} + \cdots + d_2 a_{1n},$$
$$d_1 a_{11} + \cdots + d_1 a_{1i} + d_1 a_{1(i+2)} + d_2 a_{1(i+2)} + \cdots + d_2 a_{1n}.$$

Also there are two corresponding columns of $\mathbf{X}$ with respect to each of the two components above. Choose one of column 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 column of mixing matrix $\mathbf{A}$ up to a sign.

Appendix B

We only consider noise free case here. Suppose that $(\mathbf{x}_l, \mathbf{x}_j)$ is a pair, thus $\mathbf{x}_l - \mathbf{x}_j = (d_2 - d_1)\mathbf{a}_i$.

From the model (2), $\mathbf{x}_l, \mathbf{x}_j$ can be represented as follows,

$$\mathbf{x}_l = e_{11}\mathbf{a}_1 + \cdots + e_{1i}\mathbf{a}_i + \cdots + e_{1n}\mathbf{a}_n, \qquad (\text{A}\cdot 1)$$
$$\mathbf{x}_j = e_{21}\mathbf{a}_1 + \cdots + e_{2i}\mathbf{a}_i + \cdots + e_{2n}\mathbf{a}_n, \qquad (\text{A}\cdot 2)$$

where $e_{k_1 k_2} \in \{d_1, d_2\}$, $k_1 = 1, 2,\ k_2 = 1, \cdots, n$.

Thus,

$$\begin{aligned}\mathbf{x}_l - \mathbf{x}_j &= (e_{11} - e_{21})\mathbf{a}_1 + \cdots + (e_{1n} - e_{2n})\mathbf{a}_n \\ &= (d_2 - d_1)\mathbf{a}_i. \end{aligned} \qquad (\text{A}\cdot 3)$$

There exist three possible cases:

1. $e_{1i} = d_2, e_{2i} = d_1$. From (A$\cdot$3), we have,

$$\begin{aligned}&(e_{11} - e_{21})\mathbf{a}_1 + \cdots + (e_{1(i-1)} - e_{2(i-1)})\mathbf{a}_{i-1} \\ &+(e_{1(i+1)} - e_{2(i+1)})\mathbf{a}_{i+1} + \cdots \\ &+(e_{1n} - e_{2n})\mathbf{a}_n = \mathbf{0}. \end{aligned} \qquad (\text{A}\cdot 4)$$

From the solvability condition in Theorem 1, $e_{1k_1} = e_{2k_1}, k_1 = 1, \cdots, i-1, i+1, \cdots, n$.

2. $e_{1i} = d_1, e_{2i} = d_2$. We have

$$\begin{aligned}\mathbf{x}_l - \mathbf{x}_j &= (e_{11} - e_{21})\mathbf{a}_1 + \cdots + (e_{1n} - e_{2n})\mathbf{a}_n \\ &= (d_2 - d_1)\mathbf{a}_i. \end{aligned} \qquad (\text{A}\cdot 5)$$

This implies

$$c_1\mathbf{a}_1 + \cdots + c_{i-1}\mathbf{a}_{i-1} + c_{i+1}\mathbf{a}_{i+1} + \cdots + c_n\mathbf{a}_n = 2\mathbf{a}_i, \qquad (\text{A}\cdot 6)$$

where $c_1, \cdots, c_{i-1}, c_{i+1}, \cdots, c_n \in \{1, 0, -1\}$, which is in contradiction with Assumption 2. Thus this case will not occur.

3. $e_{1i} = e_{2i}$. From (A$\cdot$3), We have

$$\begin{aligned}&(e_{11} - e_{21})\mathbf{a}_1 + \cdots + (d_1 - d_2)\mathbf{a}_i - \cdots \\ &+(e_{1n} - e_{2n})\mathbf{a}_n = \mathbf{0}. \end{aligned} \qquad (\text{A}\cdot 7)$$

From that $d_1 - d_2 \neq 0$ and the solvability condition in Theorem 1, it is impossible for this case to happen.

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

If there exists another vector assumed to be $\mathbf{x}_k$ which is a matched vector of $\mathbf{x}_l$, then there are two possible cases:

1. $\mathbf{x}_l - \mathbf{x}_k = (d_2 - d_1)\mathbf{a}_i$, which leads to $\mathbf{x}_k = \mathbf{x}_j$.
2. $\mathbf{x}_k - \mathbf{x}_l = (d_2 - d_1)\mathbf{a}_i$.

From the discussion above, we have $e_{1i} = d_1$, which is in contradiction with $e_{1i} = d_2$.

Therefore, we deduce that $\mathbf{x}_l$ has only one matched vector.

Yuanqing Li was born in Hunan Province, China, in 1966. He received the B.S. degree in applied mathematics from Wuhan University, Wuhan, China, in 1988, the M.S. degree in applied mathematics from South China Normal University, Guangzhou, China, in 1994, and the Ph.D. degree in control theory and applications from South China University of Technology, Guangzhou, China, in 1977. Presently, he is a researcher in Lab. for Advanced Brain Signal Processing, RIKEN Brain Science Institute, Japan. He is also an associate professor in South China University of Technology. His research interests include Automatic Control. Blind Signal Processing, Neural Networks and Visual Computation. He is the author or coauthor of more than 40 scientific papers in journals and conference proceedings.

Andrzej Cichocki received the M.Sc. (with honors), Ph.D., and Habilitate Doctorate (Dr. Sc.) degrees, all in electrical engineering, from Warsaw University of Technology (Poland) in 1972, 1975, and 1982, respectively. Since 1972, he has been with the Institute of Theory of Electrical Engineering and Electrical Measurements at the Warsaw University of Technology, where he became a full Professor in 1991. He is the co-author of three books: Adaptive Blind Signal and Image Processing—Learning Algorithms and Applications, Wiley, 2002, MOS Switched-Capacitor and Continuous-Time Integrated Circuits and Systems (Springer-Verlag, 1989) and Neural Networks for Optimization and Signal Processing (Teubner-Wiley, 1993/94) and more than 150 research papers. He spent at University Erlangen-Nuernberg (Germany) a few years as Alexander Humboldt Research Fellow and Guest Professor. Since 1995 he has been working in the Brain Science Institute, RIKEN (Japan). as a team leader for laboratory for Open Information Systems and currently and as a head of laboratory for Advanced Brain Signal Processing. His current research interests include neural networks, signal and image processing, especially analysis and processing of multi-sensory biomedical data. He is currently associate editor of IEEE Transaction on Neural Networks and Fluctuation and noise Letters. He was 3 years the member of the Technical Committee for Neural Network for Signal Processing and recently the member of core group who established a new IEEE Circuits and Systems Technical Committee for Blind Signal Processing. More information about his research activities can be found at http://www.bsp.brain.riken.go.jp/ .

Liqing Zhang received the B.S. degree in Mathematics from Hangzhou University, China. in 1983 and Ph.D. degree in Computer Science from Zhongshan University, China, in 1988. After graduation, he joined South China University of Technology, as a lecturer. From 1990 to 1995, he served as Associate Professor, and in 1995 was promoted to Full Professor in College of Electronic and Information Engineering, South China University of Technology. He joined the RIKEN Brain Science Institute, Japan as a Research Scientist in 1997. He is currently a full professor in Department of Computer Science and Engineering, Shanghai Jiaotong University, China. His research interests include Learning Theory, Adaptive Systems, Neuroinformatics and Computational Neuroscience. He is author or coauthor of more than 70 scientific papers in journals and conference proceedings.