DEFLATIONARY BLIND SOURCE EXTRACTION USING AN EXACT SOLUTION SUBSPACE SEARCHING SCHEME

Mingjian Zhang∗
Department of Information Technology
Hunan Police Academy
Changsha, Hunan 410138, China
mingjianzhang@163.com

Xiaohua Li†
Department of ECE
State University of New York at Binghamton
Binghamton, NY 13902, USA
xli@binghamton.edu

ABSTRACT

In this paper, we develop a new deflation-based blind source extraction (BSE) algorithm to estimate and extract source signals in a sequential fashion from the mixtures. At the beginning of this algorithm, the first source signal is estimated by a constrained optimization and its efficient updating rule. Then, the other source signals are estimated and extracted by an exact solution subspace searching-based deflation technique. The key merit is that this new technique can greatly alleviate the error accumulation problem suffered by traditional deflation methods.

Index Terms— Blind source extraction, deflation, exact solution subspace searching, error accumulation, source separation

1. INTRODUCTION

The problem of blind source extraction (BSE) [1–8] involves recovery of one or a subset of unknown source signals from their observed mixtures without prior knowledge of the mixing matrix. Unlike simultaneous blind source separation (BSS) [9, 10], the objective of BSE is to extract the source signals in a sequential fashion, i.e., one by one, rather than to recover all of them simultaneously. BSE is a well-known signal processing method with wide applications in noninvasive fetal ECG extraction [2, 3], EEG readiness potentials extraction [4], heart/lung sound signals separation [5], speech signal denoising and enhancing [6–8], etc.

BSE has some advantages over the simultaneous BSS [1], especially when the aim is just to extract a couple of desired signals from the mixtures of a large number of source signals. BSE can be considered as a generalization of principal component analysis (PCA). It can extract signals in a specified order according to some type of feature of the source signals. BSE can be designed to extract only a few source signals of interest, which is often more flexible in real applications. In general, the learning rules for BSE are simpler and have lower computational complexity.

A number of publications have addressed the problem of BSE. Liu et al. [11] applied the linear predictor method to BSE. Shi et al. [12] addressed the BSE problem in the case when the desired source signal has certain temporal structure. The traditional FastICA method tends to suffer from the divergent behavior for a mixture of Gaussian-only sources. To overcome this drawback, a criterion called KSICA was proposed in [13] for blind extraction of spatio-temporally nonstationary speech source. Sällberg et al. [14] proposed a speech BSE method with a fixed-point property that is valid for a range of sources including Gaussian signals. Leong et al. [15] generalized the traditional BSE methods to the case where the mixing is ill-conditioned and post-nonlinear. Sawada et al. [16] put forward a method for enhancing target source signals by using a two-stage method with independent component analysis (ICA) and time-frequency masking techniques. Washizawa et al. [17] proposed a method that does not need the strong assumptions such as independence or non-Gaussianity on source signals. In [18], a linear instantaneous differential fixed-point ICA (LI-DFICA) algorithm was developed for underdetermined mixtures.

In a typical sequential BSE algorithm, the following two cascaded steps are conducted. First, a source signal with certain special properties is estimated. Then, a deflation scheme is used to implicitly or explicitly remove the contribution of this source signal from the mixtures [19–24]. As pointed out in [19–21], one of the major problems of the deflation schemes is error accumulation. Estimation errors of source signals are propagated to the subsequent procedures. The accumulated errors will degrade the estimation accuracy of the source signals recovered later. In this paper, we propose an exact solution subspace searching-based deflationary blind source extraction (ESSS-DBSE) algorithm, which can greatly alleviate the error accumulation problem.

The rest of this paper is organized as follows. In Section 2, we formulate the BSE problem. In Section 3, we develop the new ESSS-DBSE algorithm. Simulations are conducted in Section 4 and conclusions are drawn in Section 5.

2. PROBLEM FORMULATION

Consider the instantaneous mixing model

$$\mathbf{x}(t) = \mathbf{A}\mathbf{s}(t) + \mathbf{n}(t)$$

(1)

where \( \mathbf{x}(t) = [x_1(t), \ldots, x_J(t)]^T \) is a vector of \( J \) observation signals, \( \mathbf{s}(t) = [s_1(t), \ldots, s_R(t)]^T \) is a vector of \( R \) source signals,
$A \in \mathbb{R}^{J \times R}$ is an unknown full column rank mixing matrix, and $n(t)$ represents additive noise. Note that $(\cdot)^T$ denotes transpose, and $\mathbb{R}$ means real domain.

We make the following assumptions for the source signals $s(t)$ and noise $n(t)$: 1) The source signals are zero-mean, spatially uncorrelated but temporally correlated; and 2) the additive noise $n(t)$ is a stationary, temporally white zero-mean random process independent of the source signals.

The covariance matrix of $x(t)$ with a time lag $\tau_i$ is defined as

$$C_i = E\{x(t)x^T(t + \tau_i)\}. \quad (2)$$

Since the noise is assumed temporally white, $C_i$ is not affected by the noise for nonzero time lags. Therefore, if $\tau_i \neq 0$, then $C_i = AD_iA^T$, where $D_i = E\{n(t)n^T(t + \tau_i)\}$ is a diagonal matrix.

Specifically,

$$C_i = AD_i\text{diag}(d_{i1}, \ldots, d_{iR})A^T \quad (3)$$

where $\text{diag}(d_{i1}, \ldots, d_{iR}) = D_i$.

When the number of the observation signals is more than the number of the source signals, i.e., $J > R$, a preprocessing step is usually applied to reduce the dimensionality of the observation signals. First, we compute the singular value decomposition (SVD) of $C_0 = E\{x(t)x^T(t)\}$ as

$$C_0 = \left[ U, U_n \right] \begin{bmatrix} \Sigma & 0 \\ 0 & \Sigma_n \end{bmatrix} \left[ V, V_n \right]^T \quad (4)$$

where $U \in \mathbb{R}^{J \times R}$ and $V \in \mathbb{R}^{J \times R}$ are columnwise orthonormal matrices whose columns correspond to $R$ principal singular values, $\Sigma = \text{diag}\{\sigma_1, \ldots, \sigma_R\}$ is a diagonal matrix of the $R$ principal singular values, and $\Sigma_n = \text{diag}\{\sigma_{R+1}, \ldots, \sigma_J\}$ is a diagonal matrix of the $(J-R)$ smallest singular values, i.e., $\sigma_1 \geq \sigma_2 \geq \ldots \geq \sigma_R > \sigma_{R+1} \geq \ldots \geq \sigma_J$. Next, with the left singular vectors contained in $U$, we perform the preprocessing of $x(t)$ as

$$\bar{x}(t) = U^T x(t). \quad (5)$$

Then, the covariance matrices $R_i = E\{\bar{x}(t)\bar{x}^T(t + \tau_i)\}$, $i = 1, \ldots, N$, can be expressed as

$$R_i = U^T C_i U. \quad (6)$$

BSS recovers the source signals by a separating matrix $W \in \mathbb{R}^{R \times R}$ such that

$$y(t) = W\bar{x}(t) = WU^T x(t) = GS(t) \quad (7)$$

has spatially uncorrelated components. BSS is considered successful when $y(t) = [y_{1}(t), \ldots, y_{R}(t)]^T$ is a permuted and scaled version of $s(t)$, which implies that the global mixture-separation matrix $G = WU^T A$ is a generalized permutation matrix of the form

$$G = PA \quad (8)$$

where $P$ is a permutation matrix and $A$ is a non-singular diagonal matrix.

Let us denote by $w$ a row of the separating matrix $W$. According to (7) and (8), $A^T W^T = e_k$, where $e_k = [0, \ldots, 0, 1, 0, \ldots, 0]^T$ is a vector with only one nonzero element $\gamma_k$ in the $k$th place. Hence,

$$R_i w^T = d_{ik}\gamma_k U^T a_k, \quad R_i w^T = d_{ik}\gamma_k U^T a_k \quad (9)$$

where $a_k$ denotes the $k$th column vector of the mixing matrix $A$.

For BSE, the extraction operation boils down to estimating a row vector $w$ at a time to extract a source signal as $y(t) = W\bar{x}(t)$.

### 3. EXACT SOLUTION SUBSPACE SEARCHING ALGORITHM

#### 3.1. Extraction Detecting Device

According to (9), the following equation holds

$$R_i w^T = \gamma R_j w^T \quad (10)$$

where $\gamma$ is a scalar. The vector $w^T$ is the generalized eigenvector of the matrix pencil $(R_i, R_j)$, and can be roughly estimated by the generalized eigenvalue decomposition (GEVD). Nevertheless, in order to obtain more accurate and non-trivial solutions, we need to use a series of covariance matrices $R_i, \ldots, R_N$ rather than only two of them.

Define $y_j = R_i w^T = [r_{ij1}, \ldots, r_{ijR}]^T w^T = [y_{ij1}, \ldots, y_{ijR}]^T$, $y_j = R_i w^T = [r_{ij1}, \ldots, r_{ijR}]^T w^T = [y_{ij1}, \ldots, y_{ijR}]^T$, where $r_{ik}$ and $r_{jk}$ are the $k$th row vectors of $R_i$ and $R_j$, respectively. According to (10), $y_i$ and $y_j$ are proportional. Therefore, we have

$$y_i y_{jt} - y_{it} y_{jt} = 0, \quad 1 \leq i < j \leq N, 1 \leq s < t \leq R. \quad (11)$$

With vector notation, (11) can be rewritten as

$$(r_{is} w^T r_{jt} - r_{it} w^T r_{js}) w^T = 0 \quad (12)$$

or

$$(r_{jt} w^T r_{is} - r_{js} w^T r_{it}) w^T = 0. \quad (13)$$

Furthermore, equations (12) and (13) can be written more concisely as

$$P w^T = 0 \quad (14)$$

where $P \in \mathbb{R}^{R \times R}$, $M = \sum_{N=1}^{R} (R-1)$, and $0$ is a zero vector. The $M$ rows of $P$ are $r_{is} w^T r_{it} - r_{it} w^T r_{is} + r_{jt} w^T r_{js} - r_{js} w^T r_{jt}, 1 \leq i < j \leq N, 1 \leq s < t \leq R$. Because (14) provides a criterion for detecting whether the extraction is achieved or not, we term it “extraction detecting device”. Note that (14) is the starting point of our new BSE algorithm that will be derived in the next two subsections.

#### 3.2. Optimization and Iterative Update Rule for First Signal Extraction

To extract a source signal from the mixtures, let us consider the constrained optimization problem

$$\min f(b) = b^T B b, \quad \text{ s.t. } ||b|| = 1$$

where $b \in \mathbb{R}^{1 \times R}$ and $|| \cdot ||$ denotes the Frobenius norm. The $M$ rows of $B$ are $r_{is} b^T r_{js} - r_{it} b^T r_{js} + r_{jt} b^T r_{js} - r_{js} b^T r_{jt}, 1 \leq i < j \leq N, 1 \leq s < t \leq R$. From (14), the solution to this constrained optimization is an estimate of $w$.

For this nonlinear optimization problem, we can use the iterative update rule [25] to search for the solution

$$b \leftarrow b + v, \quad b \leftarrow b \frac{b}{||b||} \quad (15)$$

where $v$ is the eigenvector of the matrix $B^T B$ corresponding to the eigenvalue with the smallest magnitude. Note that the matrix $B$ has to be updated as well by the vector $b$ estimated in the previous iteration. After reaching a stationary point, $b$ is an estimated row $\bar{w}_1$ of $W$ and used to extract a source signal as $\hat{s}_1(t) = b^T \bar{x}(t)$.
3.3. Exact Solution Subspace Searching-Based Deflation for the Extraction of Other Source Signals

After extracting one signal component, we can apply the deflation technique to extract other source signals from the mixtures sequentially. Let us assume that we have estimated \( l \) rows of the matrix \( \mathbf{W} \), denoted as \( \hat{w}_1, \ldots, \hat{w}_l \). With them, we have estimated \( l \) source signals, denoted as \( \hat{s}_i(t) = \hat{w}_i \mathbf{x}(t) \). Then, our objective is to find \( \hat{w}_{l+1} \) and recover \( \hat{s}_{l+1}(t) \).

To explain the deflation principle, we denote by \( \Psi_1 \) the subspace corresponding to the solutions which minimize the objective function \( b (B^T B) b^T \), and by \( \Psi_2 \) the subspace that does not contain the previously estimated row vectors \( \hat{w}_1, \ldots, \hat{w}_l \). The solution of \( \hat{w}_{l+1} \) belongs to their intersection \( \Psi_1 \cap \Psi_2 \), which we term "exact solution subspace".

To address \( \Psi_2 \), we design a penalty term \( b (uu^T) b^T \), where \( u \in \mathbb{R}^{R \times 1} \) is orthogonal to \( \hat{w}_1, \ldots, \hat{w}_l \). The application of this penalty term prevents the vector \( b \) from converging to \( \hat{w}_1, \ldots, \hat{w}_l \) and consequently fixes it to the subspace \( \Psi_2 \).

In order to search for the solutions in the subspace \( \Psi_1 \cap \Psi_2 \), which we term "exact solution subspace searching" (ESSS), we combine the objective function and the penalty term together as

\[
\lambda = \frac{b (B^T B) b^T}{b (uu^T) b^T}.
\]

As a special case, if \( b \) makes \( b (B^T B) b^T = 0 \) and \( b (uu^T) b^T \neq 0 \), then \( \lambda \) takes the minimum value of 0, and \( b \) is the required \( \hat{w}_{l+1} \). Hence, the minimization of (16) forces \( b \) to be different from \( \hat{w}_1, \ldots, \hat{w}_l \).

Considering that

\[
(B^T B) b^T = \lambda (uu^T) b^T
\]

is a sufficient condition for (16), we can limit our solution search within the generalized eigenvectors of the matrix pencil \((B^T B, uu^T)\).

This can greatly enhance convergence.

Therefore, in the deflation procedure of the sequential BSE, we first update the matrix \( B \) using the value of \( b \) estimated in the previous iteration. Then, we calculate

\[
v^* = \arg \min_{b \in \{v_1, \ldots, v_R\}} \frac{b (B^T B) b^T}{b (uu^T) b^T} \]

where \( v_1, \ldots, v_R \) are the generalized eigenvectors of the matrix pencil \((B^T B, uu^T)\). Next, we update \( b \) as

\[
b \leftarrow b + v^*, \quad b \leftarrow \frac{b}{\|b\|}.
\]

These three steps are executed iteratively until the stationary point is reached.

In order to find the vector \( u \) orthogonal to \( \hat{w}_1, \ldots, \hat{w}_l \), we first stack \( \hat{w}_1, \ldots, \hat{w}_l \) into the matrix \( \hat{W} = [\hat{w}_1^T, \ldots, \hat{w}_l^T] \). Then, we compute the eigenvalue decomposition (EVD)

\[
\hat{W} (\hat{W}^T)^T = [U_e, U_l] \begin{bmatrix} \Sigma_e & 0 \\ 0 & 0 \end{bmatrix} [U_e, U_l]^T.
\]

The vector \( u \) can be chosen from the columns of \( U_l \in \mathbb{R}^{R \times (R-l)} \).

After convergence, \( b \) is the \((l+1)\)th estimated row \( \hat{w}_{l+1} \) of \( \hat{W} \).

Consequently, the \((l+1)\)th estimated source signal can be computed as \( \hat{s}_{l+1}(t) = \hat{w}_{l+1} \mathbf{x}(t) \). This sequential extraction procedure can be continued until all the desired source signals are estimated.

It is worth noting that the estimation errors of \( \hat{w}_1, \ldots, \hat{w}_l \) affect the estimation accuracy of the vector \( u \) only. They do not affect \( b (B^T B) b^T \) explicitly. Therefore, as long as we find a vector \( b \) that minimizes \( b (B^T B) b^T \) while keeping \( b \) different from \( \hat{w}_1, \ldots, \hat{w}_l \), \( b \) is an estimate of \( \hat{w}_{l+1} \) and the estimation errors of \( \hat{w}_1, \ldots, \hat{w}_l \) are not propagated to \( \hat{w}_{l+1} \). Hence, it is possible that \( \hat{w}_{l+1} \) is as accurate as \( \hat{w}_1, \ldots, \hat{w}_l \). This means that the new deflationary BSE procedure can effectively mitigate error accumulation.

4. SIMULATIONS

In this section, we illustrate the performance of our proposed ESSS-DBSE algorithm and compare it with the SOBI [26] and FastICA [27] algorithms by simulations. We consider \( R = 3 \) speech source signals, each is 2 seconds long (see Fig. 1). They are the truncated versions of the sound signals provided by [28]. These source signals are mixed with a randomly generated \( 5 \times 3 \) mixing matrix

\[
A = \begin{bmatrix} -1.6877 & 0.8435 & -0.3982 \\ -0.2161 & 0.9629 & 1.0942 \\ -0.7075 & 0.6267 & 1.8387 \\ 0.7050 & -0.2705 & 0.3615 \\ -0.6629 & -0.9120 & -0.8982 \end{bmatrix}.
\]

Fig. 2 shows the observed mixtures of the speech source signals.
4.1. Performance Comparison

In simulations, noise is added to introduce signal-to-noise ratios (SNRs) varying from -5 dB to 30 dB. The number of covariance matrices is $N = 10$. The vector $b$ is initialized with a row of $W$ which is roughly estimated by the GEVD of the matrix pencil $(R_1, R_2)$. As performance measure, we use the performance index [1]

$$\text{PI} = \frac{1}{R(R-1)} \sum_{i=1}^{R} \left\{ \left( \sum_{k=1}^{R} \frac{|\hat{g}_{ik}|}{\max_j |\hat{g}_{ij}|} - 1 \right) + \left( \sum_{k=1}^{R} \frac{|\hat{g}_{ki}|}{\max_j |\hat{g}_{ij}|} - 1 \right) \right\}$$

(21)

where $\hat{g}_{ij}$ is the $(i,j)$-element of the estimated global matrix $G = WU^T A$. The PI measures to what extent the estimated global matrix is close to a generalized permutation matrix. Obviously, the smaller the value of PI, the better the separation performance.

For FastICA, we utilized the MATLAB code from the website http://www.cis.hut.fi/projects/ica/fastica/ and used the deflation type to extract source signals sequentially rather than simultaneously. The ESSS-DBSE algorithm and the SOBI algorithm are simulated with the same set of covariance matrices $\{C_1, \ldots, C_N\}$. The SOBI algorithm is based on the joint diagonalization technique and recovers source signals simultaneously. We used 100 independent runs to calculate the average PI. Simulation results are shown in Fig. 3.

It can be seen that, with SNR $\in [-5, 10]$ dB, the performance of the proposed ESSS-DBSE algorithm is close to that of SOBI. However, with SNR $\in [15, 30]$ dB, the performance of ESSS-DBSE is much better than that of SOBI. In contrast, the performance of FastICA is much worse, especially when the SNR is not too high. This means that the ESSS-DBSE and SOBI algorithms are more robust to noise. In general, the proposed ESSS-DBSE algorithm achieves better performance than SOBI and FastICA.

4.2. Error Accumulation and Convergence of ESSS-DBSE

To evaluate the estimation accuracy of the $i$th extracted source signal, we use the accuracy index

$$\epsilon_i = \sum_{k=1}^{R} \frac{|\hat{g}_{ik}|}{\max_j |\hat{g}_{ij}|} - 1.$$

The smaller the value of $\epsilon_i$, the better the $i$th extracted source signal in the sequential extraction procedure.

We first compare $\epsilon_i$, $i = 1, \ldots, R$, to check whether the proposed ESSS-DBSE algorithm suffers from error accumulation. Let $O_k$, $k = 1, \ldots, 6$, denote the orders $[1 \ 2 \ 3]$, $[2 \ 1 \ 3]$, $[3 \ 1 \ 2]$, $[3 \ 2 \ 1]$, and $[3 \ 1 \ 2]$, and $[3 \ 2 \ 1]$, respectively. For instance, if $\epsilon_3 < \epsilon_1 < \epsilon_2$, then this case corresponds to the order $O_3$. Based on $\epsilon_i$, we count the number of $O_k$, $k = 1, \ldots, 6$, over 100 independent trials. The results are shown in Table 1. It can be seen that the condition $\epsilon_1 < \epsilon_j$ for any $1 \leq i < j \leq R$ does not hold. In other words, the source signals extracted later do not necessarily have worse accuracy than those extracted earlier. This means that the proposed ESSS-DBSE algorithm can effectively mitigate error accumulation during the deflation procedure.

<table>
<thead>
<tr>
<th>SNR (dB)</th>
<th>PI (dB)</th>
</tr>
</thead>
<tbody>
<tr>
<td>-5</td>
<td>-30</td>
</tr>
<tr>
<td>0</td>
<td>-25</td>
</tr>
<tr>
<td>5</td>
<td>-20</td>
</tr>
<tr>
<td>10</td>
<td>-15</td>
</tr>
<tr>
<td>15</td>
<td>-10</td>
</tr>
<tr>
<td>20</td>
<td>-5</td>
</tr>
<tr>
<td>25</td>
<td>0</td>
</tr>
<tr>
<td>30</td>
<td>5</td>
</tr>
</tbody>
</table>

Table 1. Number of Orders in 100 Independent Runs

<table>
<thead>
<tr>
<th>Orders</th>
<th>$O_1$</th>
<th>$O_2$</th>
<th>$O_3$</th>
<th>$O_4$</th>
<th>$O_5$</th>
<th>$O_6$</th>
</tr>
</thead>
<tbody>
<tr>
<td>SNR=−5dB</td>
<td>8</td>
<td>28</td>
<td>13</td>
<td>13</td>
<td>22</td>
<td>16</td>
</tr>
<tr>
<td>SNR=0dB</td>
<td>21</td>
<td>23</td>
<td>15</td>
<td>11</td>
<td>14</td>
<td>16</td>
</tr>
<tr>
<td>SNR=5dB</td>
<td>19</td>
<td>19</td>
<td>17</td>
<td>18</td>
<td>18</td>
<td></td>
</tr>
<tr>
<td>SNR=10dB</td>
<td>35</td>
<td>26</td>
<td>7</td>
<td>7</td>
<td>13</td>
<td>12</td>
</tr>
<tr>
<td>SNR=15dB</td>
<td>20</td>
<td>30</td>
<td>4</td>
<td>5</td>
<td>29</td>
<td>12</td>
</tr>
<tr>
<td>SNR=20dB</td>
<td>14</td>
<td>34</td>
<td>7</td>
<td>1</td>
<td>37</td>
<td>7</td>
</tr>
<tr>
<td>SNR=25dB</td>
<td>5</td>
<td>32</td>
<td>6</td>
<td>0</td>
<td>56</td>
<td>1</td>
</tr>
<tr>
<td>SNR=30dB</td>
<td>2</td>
<td>12</td>
<td>7</td>
<td>0</td>
<td>79</td>
<td>0</td>
</tr>
</tbody>
</table>

When SNR is high, the performance difference between ESSS-DBSE and FastICA mainly stems from their abilities to refrain from error accumulation during the deflation procedure. The numbers of $O_k$ of FastICA are 24, 0, 12, 58, 6, and 0, respectively, over 100 independent trials for SNR=30dB. It can be seen that ESSS-DBSE has stronger ability to refrain from error accumulation than FastICA.

Next, we evaluate the convergence speed of the iterative update rule for the proposed ESSS-DBSE algorithm. We count the number of iterations and average it over 100 independent trials. Table 2 shows the average number of iterations. From Table 2, it can be seen that only a small number of iterations are required for the convergence of the ESSS-DBSE algorithm.

<table>
<thead>
<tr>
<th>SNR (dB)</th>
<th>Iterations</th>
</tr>
</thead>
<tbody>
<tr>
<td>-5</td>
<td>7</td>
</tr>
<tr>
<td>0</td>
<td>6</td>
</tr>
<tr>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>10</td>
<td>4</td>
</tr>
<tr>
<td>15</td>
<td>4</td>
</tr>
<tr>
<td>20</td>
<td>4</td>
</tr>
<tr>
<td>25</td>
<td>4</td>
</tr>
<tr>
<td>30</td>
<td>4</td>
</tr>
</tbody>
</table>

Table 2. Number of Iterations Required for Convergence

5. CONCLUSION

We developed a new algorithm for sequential blind source extraction from instantaneous mixtures based on the exact solution subspace searching technique. The merits of the proposed algorithm include the ability to refrain from error accumulation during the deflation procedure and the high convergence speed. Simulations are conducted to demonstrate the superior performance of the proposed algorithm.
6. REFERENCES


