Using Joint Generalized Eigenvectors of a Set of Covariance Matrix Pencils for Deflationary Blind Source Extraction

Mingjian Zhang, *Member, IEEE*, Xiaohua Li, *Senior Member, IEEE*, and Jun Peng, *Member, IEEE*

***Abstract*—In this paper, we develop a new deflationary blind source extraction (BSE) algorithm that extracts source signals in a sequential fashion via the joint generalized eigenvectors of a set of covariance matrix pencils. The new concept of joint generalized eigenvector is defined. We prove that these vectors can be made unique and identical to the source extraction vectors with properly selected matrix pencils. To resolve the open problem of estimating joint generalized eigenvectors, we develop an approach based on the deflation operation and the proportional property of the joint generalized eigenvectors. Specifically, with the proportional property, we show that the estimation problem can be formulated as an optimization involving a quadratic cost function and a unit-rank matrix constraint. An efficient iterative algorithm is then developed by applying the gradient search, matrix shrinkage, deflation, and symmetry-preserving vectorization techniques. This algorithm estimates the joint generalized eigenvectors and conducts BSE sequentially. Its computational complexity and convergence are analyzed. Simulations demonstrate that this algorithm outperforms many typical BSE or blind source separation algorithms. In particular, the new algorithm is more robust to both heavy noise and ill-conditioned mixing matrices.**

***Index Terms*—Source separation, blind source extraction, joint generalized eigenvector, matrix shrinkage, matrix rank minimization, gradient search algorithm.**

I. Introduction

BLIND source extraction (BSE) involves recovery of one or a subset of unknown source signals from the observed mixtures without knowing the mixing matrix [1]. BSE has found wide applications in many fields, such as noninvasive fetal ECG extraction [2], [3], heart/lung sound signals separation [4], EEG readiness potentials extraction [5], speech signal denoising and enhancing [6]–[8], etc.

Unlike blind source separation (BSS) that recovers all source signals simultaneously [9]–[11], BSE extracts source signals one by one in a sequential fashion. A typical BSE algorithm conducts the following two steps iteratively: i) recover a source signal, and ii) apply a deflation procedure to remove this source signal from the mixture [12]–[20]. Sequential BSE has a number of advantages over simultaneous BSS [1]. It is more efficient in computation and more flexible in practical applications.

BSE has attracted a lot of research attention during the recent decade. Liu *et al.* [21] applied linear prediction techniques for BSE. Shi *et al.* [22] addressed it by exploiting the temporal structure of the desired source signals. A criterion called KSICA was proposed in [23] for the blind extraction of spatio-temporally nonstationary speech sources. Leong *et al.* [24] generalized BSE to the case where the mixing function is ill-conditioned and post-nonlinear. Sawada *et al.* [25] put forward a method to enhance target source signals via the independent component analysis (ICA) and time-frequency masking techniques. A scheme for extracting source signals with the highest auto-correlation was proposed in [26]. The method of Washizawa *et al.* [27] conducted BSE without the need of strong source assumptions such as independence and non-Gaussianity. In [28], a linear instantaneous differential fixed-point ICA (LI-DFICA) algorithm was developed to extract sources from under-determined mixtures.

On the other hand, although many algorithms have been developed, considering the wide application of BSE in a variety of potential fields, more efforts are still needed to develop new algorithms with enhanced performance and robustness.

In this paper, we investigate a completely new BSE approach based on the joint generalized eigenvectors of a set of matrix pencils. The matrix pencils are formed by the covariance matrices of the observed mixtures. In order to make the theoretical development easier to follow, we outline the major steps of our deductions in Fig. 1. Specifically, first, we define the *joint generalized eigenvectors* and prove their equivalence to *source extraction vectors*. Then, we formulate the estimation of these vectors into a quadratic optimization over a unit-rank matrix. Finally, this optimization is solved by an efficient iterative algorithm which gives both the estimated vectors and the extracted sources.

Manuscript received April 16, 2017; revised September 27, 2017 and January 2, 2018; accepted January 4, 2017. Date of publication January 11, 2018; date of current version April 19, 2018. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Eleftherios Kofidis. This work was supported in part by the US National Science Foundation under Grant CNS-1443885, in part by the National Natural Science Foundation of China under Grants 61379111, 61471169, and 61672537, in part by the National Social Science Foundation of China under Grant 15BYY097, and in part by the Humanities and Social Science Project of Ministry of Education of China under Grant 13YJCZH250. *(Corresponding author: Xiaohua Li.)*

M. Zhang is with the School of Information Science and Engineering, Central South University, Changsha, Hunan, China, and also with Hunan Police Academy, Changsha 410138, China (e-mail: mingjianzhang@163.com).

X. Li is with the Department of Electrical and Computer Engineering, State University of New York at Binghamton, Binghamton, NY 13902 USA (e-mail: xli@binghamton.edu).

J. Peng is with the School of Information Science and Engineering, Central South University, Changsha 410083, China (e-mail: pengj@csu.edu.cn).

This paper has supplementary downloadable material available at http://ieeexplore.ieee.org, provided by the author. The material includes the MATLAB source code for the new algorithm.

Digital Object Identifier 10.1109/TSP.2018.2791980

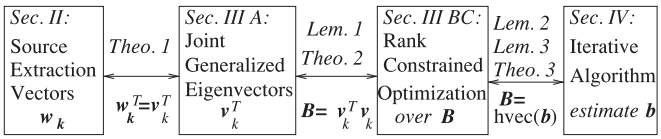

Fig. 1. Outline of the major work in this paper, including the role of theorems (Theo.) and lemmas (Lem.).

Our major contributions in this paper include:

- We define the new concept of joint generalized eigenvectors and prove their uniqueness and equivalence to source extraction vectors in BSE applications.
- We show that the joint generalized eigenvectors can be estimated via a rank constrained optimization. This gives a practical way to solve the open problem regarding how to find these vectors.
- We develop an efficient iterative algorithm to estimate the joint generalized eigenvectors and to use them for BSE. This algorithm uses gradient search to minimize a quadratic cost function, applies matrix shrinkage to meet the unit-rank constraint, exploits symmetry-preserving vectorization to reduce complexity, and adopts deflation operation to estimate all the vectors and source signals sequentially.
- The convergence of the new algorithm is proved. Extensive simulations are conducted to demonstrate its superior performance and robustness.

The rest of the paper is organized as follows. In Section II, we formulate the BSE problem. In Section III, we define the joint generalized eigenvectors and investigate their estimations. In Section IV, the new algorithm is developed. Simulations are conducted in Section V and conclusions are given in Section VI.

Notations: The superscript $(\cdot)^T$ denotes transpose, $\mathbb{R}$ denotes real domain, $E\{\cdot\}$ denotes expectation, $\mathrm{tr}(\cdot)$ denotes the trace of a matrix, $\|\cdot\|_F$ denotes the Frobenius norm, and $\|\mathbf{x}\|$ denotes the Euclidean norm of a vector $\mathbf{x}$. The operator norm of the matrix $\mathbf{X}$ is defined as $\|\mathbf{X}\| \triangleq \sigma_1$, where σ_1 is the largest singular value of $\mathbf{X}$. In addition, $\mathbf{e}_k = [0, \ldots, 1, \ldots, 0]^T$ denotes the unit vector with only one nonzero element (which is 1) in the kth place, $\mathrm{diag}[\mathbf{x}]$ (or $\mathrm{diag}[x_1, \ldots, x_R]$) is a diagonal matrix with the elements of $\mathbf{x}$ (or x_i) in diagonal entries, and the function $(x)^+ = \max\{0, x\}$.

II. Problem Formulation

Consider the instantaneous mixing model

$$\mathbf{x}(t) = \mathbf{A}\mathbf{s}(t) + \mathbf{n}(t) \tag{1}$$

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, $\mathbf{A} \in \mathbb{R}^{J \times R}$ is an unknown full column rank mixing matrix, and $\mathbf{n}(t)$ represents additive noise. We assume that the source signals are real, zero-mean, spatially uncorrelated but temporally correlated. The additive noise $\mathbf{n}(t)$ is a real, stationary, temporally white zero-mean random process independent of the source signals.

The covariance matrix of $\mathbf{x}(t)$ with a time lag τ_i is defined as $\mathbf{C}_{x,i} \triangleq E\{\mathbf{x}(t)\mathbf{x}^T(t+\tau_i)\}$. Without any loss of generality, we consider discrete-time signals, where both t and τ_i are integers. Since the noise is assumed temporally white, if $\tau_i \neq 0$, then

$$\mathbf{C}_{x,i} = \mathbf{A}\mathbf{D}_i\mathbf{A}^T \tag{2}$$

where $\mathbf{D}_i \triangleq E\{\mathbf{s}(t)\mathbf{s}^T(t+\tau_i)\} = \mathrm{diag}[d_{i1}, \ldots, d_{iR}]$.

When the number of observation signals is more than the number of source signals, i.e., $J > R$, a standard dimension-reduction procedure is usually applied to reduce the signal dimensions [1], [29]. Specifically, first, we compute the eigenvalue decomposition (EVD) of $\mathbf{C}_{x,0} = E\{\mathbf{x}(t)\mathbf{x}^T(t)\}$, which gives

$$\mathbf{C}_{x,0} = \begin{bmatrix} \mathbf{U}_s & \mathbf{U}_n \end{bmatrix} \begin{bmatrix} \mathbf{\Sigma}_s & \mathbf{0} \\ \mathbf{0} & \mathbf{\Sigma}_n \end{bmatrix} \begin{bmatrix} \mathbf{U}_s & \mathbf{U}_n \end{bmatrix}^T, \tag{3}$$

where $\mathbf{U}_s \in \mathbb{R}^{J \times R}$ and $\mathbf{U}_n \in \mathbb{R}^{J \times (J-R)}$ are columnwise orthonormal matrices, $\mathbf{\Sigma}_s = \mathrm{diag}[\lambda_{x,1}, \ldots, \lambda_{x,R}]$ is a diagonal matrix of the R principal eigenvalues, and $\mathbf{\Sigma}_n = \mathrm{diag}[\lambda_{x,R+1}, \ldots, \lambda_{x,J}]$ is a diagonal matrix of the $(J-R)$ noise eigenvalues. Then, with the average $\bar{\lambda}_{x,n} = \frac{1}{J-R}\sum_{i=R+1}^{J} \lambda_{x,i}$, we can use the matrix $\mathbf{U} = (\mathrm{diag}[\sqrt{\lambda_{x,1} - \bar{\lambda}_{x,n}}, \ldots, \sqrt{\lambda_{x,R} - \bar{\lambda}_{x,n}}])^{-1}\mathbf{U}_s^T$ to whiten (and reduce the dimension of) the signal $\mathbf{x}(t)$ as

$$\bar{\mathbf{x}}(t) = \mathbf{U}\mathbf{x}(t). \tag{4}$$

For the signal $\bar{\mathbf{x}}(t)$, the new covariance matrix is

$$\mathbf{R}_i \triangleq E\{\bar{\mathbf{x}}(t)\bar{\mathbf{x}}^T(t+\tau_i)\} = \mathbf{U}\mathbf{A}\mathbf{D}_i\mathbf{A}^T\mathbf{U}^T. \tag{5}$$

In source separation, we recover the source signals by a separating matrix $\mathbf{W} \in \mathbb{R}^{R \times R}$, which gives

$$\mathbf{y}(t) = \mathbf{W}\bar{\mathbf{x}}(t). \tag{6}$$

From (4) and (1), it can be seen that

$$\mathbf{y}(t) = \mathbf{W}\mathbf{U}\mathbf{x}(t) = \mathbf{G}\mathbf{s}(t) + \mathbf{W}\mathbf{U}\mathbf{n}(t) \tag{7}$$

where $\mathbf{G} = \mathbf{W}\mathbf{U}\mathbf{A}$. Source separation is considered successful when $\mathbf{y}(t) = [y_1(t), \ldots, y_R(t)]^T$ is a permuted and scaled version of $\mathbf{s}(t)$ plus noise, which implies that $\mathbf{G}$ is a generalized permutation matrix, i.e.,

$$\mathbf{G} = \mathbf{P}\mathbf{\Lambda} \tag{8}$$

with a permutation matrix $\mathbf{P}$ and a non-singular diagonal matrix $\mathbf{\Lambda} = \mathrm{diag}[\gamma_1, \ldots, \gamma_R]$.

According to (5)–(8), the matrix $\mathbf{W}\mathbf{R}_i\mathbf{W}^T$ is diagonal. A popular approach for source separation is thus to find an orthogonal matrix $\mathbf{W}$ that simultaneously diagonalizes a set of N covariance matrices $\mathbf{R}_i$, $i = 1, \ldots, N$, and this task can be conducted via approximate joint diagonalization (AJD) techniques [29]–[48]. In this paper, instead of conducting AJD directly, we show that a new concept of joint generalized eigenvector can potentially lead to more robust BSE algorithms.

III. Joint Generalized Eigenvectors and Their Estimation

A. Joint Generalized Eigenvector and Source Extraction Vector

Assume that there is an orthonormal matrix $\mathbf{W}$ that can successfully separate all the R sources. Denote by $\mathbf{w}_k$ the row of $\mathbf{W}$ that extracts the kth source signal, $k = 1, \ldots, R$. Then $\mathbf{w}_k^T$ is called the *source extraction vector*. According to (7) and (8), we have $\mathbf{w}_k \mathbf{U}\mathbf{A} = \gamma_k \mathbf{e}_k^T$.

Among the N covariance matrices $\mathbf{R}_i$, $i = 1, \ldots, N$, consider an arbitrary pair $\mathbf{R}_i$ and $\mathbf{R}_j$, $i \neq j$. From (5) we have

$$\mathbf{R}_i \mathbf{w}_k^T = d_{ik}\gamma_k \mathbf{U}\mathbf{a}_k, \quad \mathbf{R}_j \mathbf{w}_k^T = d_{jk}\gamma_k \mathbf{U}\mathbf{a}_k, \tag{9}$$

where $\mathbf{a}_k$ denotes the kth column of the mixing matrix $\mathbf{A}$. The two equations in (9) lead to

$$\mathbf{R}_i \mathbf{w}_k^T = \lambda_{i,j} \mathbf{R}_j \mathbf{w}_k^T, \tag{10}$$

where $\lambda_{i,j}$ is a scalar. Hence, the R vectors $\mathbf{w}_k^T$ are all generalized eigenvectors of the matrix pencil $(\mathbf{R}_i, \mathbf{R}_j)$.

Blind source separation by the generalized eigenvalue decomposition (GEVD) of a single matrix pencil was discussed in Chapter 4 of [1]. The experience indicates that, while we can obtain R generalized eigenvectors from the GEVD of a single matrix pencil $(\mathbf{R}_i, \mathbf{R}_j)$, these R vectors may not be exactly the R source extraction vectors $\mathbf{w}_k^T$. For example, if there are repetitive generalized eigenvalues in the matrix pencil $(\mathbf{R}_i, \mathbf{R}_j)$, then GEVD may give linear combinations of $\mathbf{w}_k^T$ rather than $\mathbf{w}_k^T$ themselves. Obviously, linear combinations of $\mathbf{w}_k^T$ cannot be used as source extraction vectors.

In this paper, we will show that we can resolve this problem by considering the joint generalized eigenvectors of multiple matrix pencils $(\mathbf{R}_i, \mathbf{R}_j)$, $i \neq j$.

Definition 1: The *joint generalized eigenvector* of a set of matrix pencils $(\mathbf{R}_i, \mathbf{R}_j)$, $i \neq j$, is the vector $\mathbf{v}^T$ that satisfies $\mathbf{R}_i \mathbf{v}^T = \lambda_{i,j} \mathbf{R}_j \mathbf{v}^T$ for all $i \neq j$, where $\lambda_{i,j}$ are scalars and $\|\mathbf{v}^T\| = 1$.

In this paper, we consider the joint generalized eigenvectors of the set of matrix pencils $(\mathbf{R}_i, \mathbf{R}_j)$, $(i,j) \in \mathcal{I} \subseteq \{(i,j) : 1 \leq i < j \leq N\}$. Specifically, the set $\mathcal{I}$ has M elements, which are selected from the full set of $N(N-1)/2$ matrix pencil indices $\{(i,j) : 1 \leq i < j \leq N\}$.

Theorem 1: The set of matrix pencils $(\mathbf{R}_i, \mathbf{R}_j)$, $(i,j) \in \mathcal{I}$, have R unique joint generalized eigenvectors $\mathbf{v}_k^T$ that are equal to the R source extraction vectors $\mathbf{w}_k^T$ if and only if

$$\forall 1 \leq m \neq n \leq R, \ \exists (i,j) \in \mathcal{I}, \quad \text{such that} \quad d_{im} d_{jn} - d_{in} d_{jm} \neq 0. \tag{11}$$

Proof: The proof is given in Appendix A. ■

Theorem 1 shows that we can make the joint generalized eigenvectors unique and identical to the source extraction vectors by selecting appropriately the set $\mathcal{I}$ for (11) to be satisfied. For example, if (11) can be satisfied with $\mathcal{I} = \{(1,2)\}$, then we just need to find the R generalized eigenvectors of the matrix pencil $(\mathbf{R}_1, \mathbf{R}_2)$ and use them as source extraction vectors. Otherwise, we have to use more matrix pencils, such as $\mathcal{I} = \{(1,2), (1,3)\}$, etc, and estimate their joint generalized eigenvectors.

For source signals with rich temporal correlations, we need only a small number of matrix pencils, which means $M \ll N(N-1)/2$. On the other hand, using more matrix pencils can enhance the accuracy of estimating $\mathbf{w}_k^T$ and $\mathbf{v}_k^T$ from noisy signals. Note that we will use $\mathbf{v}_k^T$ and $\mathbf{w}_k^T$ interchangeably in sequel considering their equivalence.

B. Properties of Joint Generalized Eigenvectors

The estimation of joint generalized eigenvectors is a nontrivial open problem. Conventional GEVD works on a single matrix pencil only. AJD or other simultaneous matrix diagonalization algorithms such as [49] may not be desirable because they target toward the (more stringent) multiple-matrix diagonalization problem. Before developing practical estimation algorithms to fill this gap, we show some useful properties of joint generalized eigenvectors first in this subsection.

Consider a joint generalized eigenvector $\mathbf{v}^T$. Define $\mathbf{u}_i \triangleq \mathbf{R}_i \mathbf{v}^T$. Then from (10) we have $\mathbf{u}_i = \lambda_{i,j} \mathbf{u}_j$. In other words, $\mathbf{u}_i$ and $\mathbf{u}_j$ are proportional, which we call the *proportional property* of joint generalized eigenvectors.

Lemma 1: For two non-zero vectors $\mathbf{u}_i \in \mathbb{R}^{R \times 1}$ and $\mathbf{u}_j \in \mathbb{R}^{R \times 1}$, let u_{is} and u_{js} be their sth elements, respectively, where $1 \leq s \leq R$. Then $\mathbf{u}_i$ and $\mathbf{u}_j$ are proportional if and only if

$$u_{is} u_{jt} - u_{it} u_{js} = 0, \quad 1 \leq s < t \leq R. \tag{12}$$

Proof: See Appendix B. ■

Let $\mathbf{r}_{is}$ be the sth row of $\mathbf{R}_i$. We have $u_{is} = \mathbf{r}_{is} \mathbf{v}^T$, and the new proportional property (12) leads to

$$(\mathbf{r}_{is} \mathbf{v}^T)(\mathbf{r}_{jt} \mathbf{v}^T) - (\mathbf{r}_{it} \mathbf{v}^T)(\mathbf{r}_{js} \mathbf{v}^T) = 0, \tag{13}$$

which can be rewritten as

$$\mathbf{v}(\mathbf{r}_{is}^T \mathbf{r}_{jt} - \mathbf{r}_{it}^T \mathbf{r}_{js}) \mathbf{v}^T = 0, \ \ (i,j) \in \mathcal{I}, \ 1 \leq s < t \leq R. \tag{14}$$

Lemma 1 indicates that (14) is equivalent to (10).

For each matrix pencil $(\mathbf{R}_i, \mathbf{R}_j)$, $(i,j) \in \mathcal{I}$, we can define a list of $R(R-1)/2$ symmetric matrices

$$\mathbf{R}_{st}^{ij} \triangleq (\mathbf{r}_{is}^T \mathbf{r}_{jt} - \mathbf{r}_{it}^T \mathbf{r}_{js}) + (\mathbf{r}_{jt}^T \mathbf{r}_{is} - \mathbf{r}_{js}^T \mathbf{r}_{it}), \tag{15}$$

where $1 \leq s < t \leq R$. Then, from (14) we have

$$\mathbf{v} \mathbf{R}_{st}^{ij} \mathbf{v}^T = 0, \quad (i,j) \in \mathcal{I}, \ 1 \leq s < t \leq R, \tag{16}$$

which we call the *quadratic property* of joint generalized eigenvectors. This property serves as our starting point in developing new algorithms for the estimation of joint generalized eigenvectors.

C. Rank Constrained Optimization

Consider the problem of estimating the vector $\mathbf{v}^T$ based purely on (16). First, while $\mathbf{v}^T$ is the eigenvector associated with the zero eigenvalue of $\mathbf{R}_{st}^{ij}$, it is nontrivial to estimate a

common set of R vectors $\mathbf{v}^T$ for all the matrices. Second, because the matrices $\mathbf{R}_{st}^{ij}$ may not be semi-definite, $\mathbf{v}^T$ cannot be obtained in closed-form by forcing the derivative of $\mathbf{v}\mathbf{R}_{st}^{ij}\mathbf{v}^T$ equal to zero. Instead, we have to minimize $(\mathbf{v}\mathbf{R}_{st}^{ij}\mathbf{v}^T)^2$, which is no longer quadratic. A standard technique is to relax $\mathbf{v}^T\mathbf{v}$ into a symmetric unit-rank matrix $\mathbf{B}$ [50]. Specifically, since (16) can be written as

$$\text{tr}(\mathbf{R}_{st}^{ij}\mathbf{v}^T\mathbf{v}) = 0, \tag{17}$$

we relax it with an $R \times R$ symmetric matrix $\mathbf{B}$ into

$$\text{tr}(\mathbf{R}_{st}^{ij}\mathbf{B}) = 0, \quad \text{s.t.} \quad \text{rank}(\mathbf{B}) = 1. \tag{18}$$

Considering all the matrices $\mathbf{R}_{st}^{ij}$, we have the optimization

$$\min \ f(\mathbf{B}), \quad \text{s.t.} \quad \text{rank}(\mathbf{B}) = 1, \tag{19}$$

where $f(\mathbf{B}) = \sum_{(i,j)\in\mathcal{I}, 1\le s<t\le R}(\text{tr}(\mathbf{R}_{st}^{ij}\mathbf{B}))^2$.

Theorem 2: The optimal solution to (19) is $\mathbf{B}^* = \mathbf{w}_k^T\mathbf{w}_k$ where $\mathbf{w}_k^T$ is a joint generalized eigenvector (and is also a source extraction vector).

Proof: See Appendix C. ■

The exploitation of the new concept of joint generalized eigenvectors and the estimation of these vectors via the rank constrained optimization (19) make our approach both unique and advantageous over many existing BSE algorithms, including those that also use quadratic cost functions, such as [51]–[57]. The cost function $f(\mathbf{B})$ is quadratic, and the domain set $\{\mathbf{B} : \text{tr}(\mathbf{R}_{st}^{ij}\mathbf{B}) = 0, (i,j) \in \mathcal{I}, 1 \le s < t \le R\}$ is a convex cone over non-negative definite matrices. Such properties can potentially enhance our approach's accuracy, efficiency and convergence.

Furthermore, (19) is a variation of the matrix rank minimization problem [50] that has attracted a lot of research attention recently and has found wide applications in quadratic optimization, signal processing, system identification, etc. Many algorithms have been developed for this problem [58]–[66], which can be adapted for our purpose. In particular, the FPCA (fixed point continuation with approximate singular value decomposition) Algorithm of [58] is powerful, fast and robust, inspired by which we develop a new approach in sequel that combines the gradient search and matrix shrinkage techniques to solve (19).

As basic ideas, first, we apply gradient search to minimize the quadratic cost function $f(\mathbf{B})$. From (19), the gradient can be found as

$$\nabla f(\mathbf{B}) \triangleq \frac{\partial f(\mathbf{B})}{\partial \mathbf{B}} = \sum_{(i,j)\in\mathcal{I}, 1\le s<t\le R} 2\text{tr}(\mathbf{R}_{st}^{ij}\mathbf{B})\mathbf{R}_{st}^{ij}, \tag{20}$$

with which we can derive the updating rule

$$\mathbf{B}^{(\ell)} = \mathbf{B}^{(\ell-1)} - \alpha\nabla f(\mathbf{B}^{(\ell-1)}), \quad \ell = 1, 2, \ldots, \tag{21}$$

where α is an appropriate step size.

Next, we apply the matrix shrinkage operator $S_v(\mathbf{B}^{(\ell)})$ to satisfy the unit-rank constraint [58]. Let the singular value decomposition (SVD) of $\mathbf{B}^{(\ell)}$ be $\mathbf{B}^{(\ell)} = \mathbf{U}_B\text{diag}[\boldsymbol{\sigma}]\mathbf{V}_B^T$ where $\boldsymbol{\sigma} = [\sigma_1, \ldots, \sigma_R]$, and $\mathbf{U}_B, \mathbf{V}_B$ are orthonormal matrices. Then

$$S_v(\mathbf{B}^{(\ell)}) \triangleq \mathbf{U}_B\text{diag}[s_v(\boldsymbol{\sigma})]\mathbf{V}_B^T, \tag{22}$$

where the vector shrinkage operator $s_v(\boldsymbol{\sigma})$ is defined as

$$s_v(\boldsymbol{\sigma}) \triangleq [(\sigma_1 - v)^+, \ldots, (\sigma_R - v)^+] \tag{23}$$

with appropriate threshold $v > 0$.

However, direct implementation of the gradient search (21) is not desirable. Each iteration needs the calculation of $MR(R-1)/2$ matrices $\mathbf{R}_{st}^{ij}$, which has extremely high computational complexity. The symmetry of $\mathbf{B}$ is neither exploited nor maintained explicitly. The matrix $\mathbf{B}^{(\ell)}$ may gradually lose symmetry during updating due to the accumulation of calculation errors, which certainly degrades accuracy. To address these issues, in next section we apply symmetry-preserving vectorization and deflation techniques to implement a more efficient algorithm.

IV. Efficient Algorithm Implementation

A. Re-Formulation via Symmetry-Preserving Vectorization

It is well known that $\text{tr}(\mathbf{R}_{st}^{ij}\mathbf{B}) = (\text{vec}(\mathbf{R}_{st}^{ij}))^T\text{vec}(\mathbf{B})$, where $\text{vec}(\cdot)$ is the vectorization operator that stacks the columns of a matrix into a column vector. Considering that $\mathbf{R}_{st}^{ij}$ and $\mathbf{B}$ are both symmetric matrices, it is more desirable to remove the identical elements so as to reduce the vector dimension. For this purpose, we define three new vectorization operators $\text{svec}(\cdot)$, $\text{hvec}(\cdot)$ and $\text{unhvec}(\cdot)$. The first two operators stack an $R \times R$ symmetric matrix into an $\frac{R(R+1)}{2}$ dimensional vector, whereas the last one is the inverse operation of $\text{hvec}(\cdot)$. In addition, $\text{svec}(\cdot)$ adds the repeated elements together, while $\text{hvec}(\cdot)$ does not.

Definition 2. Symmetry-preserving vectorization operators: 1) $\text{svec}(\cdot)$ is defined as $\mathbf{b} = \text{svec}(\mathbf{B})$, where the elements are $(\mathbf{b})_{(i-1)(R-i/2)+j} = \beta_{ij}(\mathbf{B})_{ij}, 1 \le i \le j \le R, \beta_{ij} = 1$ for $i = j$, and $\beta_{ij} = 2$ for $i < j$. 2) $\text{hvec}(\cdot)$ is defined as $\mathbf{b} = \text{hvec}(\mathbf{B})$, where $(\mathbf{b})_{(i-1)(R-i/2)+j} = (\mathbf{B})_{ij}$, $1 \le i \le j \le R$. 3) $\text{unhvec}(\cdot)$ is defined as $\mathbf{B} = \text{unhvec}(\mathbf{b})$, where $(\mathbf{B})_{ij} = (\mathbf{B})_{ji} = (\mathbf{b})_{(i-1)(R-i/2)+j}$.

With the new operators, we can rewrite $f(\mathbf{B})$ of (19) as

$$\begin{aligned} f(\mathbf{b}) &= \sum_{(i,j)\in\mathcal{I}, 1\le s<t\le R} ((\text{svec}(\mathbf{R}_{st}^{ij}))^T\mathbf{b})^2 \\ &= \sum_{(i,j)\in\mathcal{I}, 1\le s<t\le R} \mathbf{b}^T\text{svec}(\mathbf{R}_{st}^{ij})(\text{svec}(\mathbf{R}_{st}^{ij}))^T\mathbf{b}, \end{aligned} \tag{24}$$

where $\mathbf{b} = \text{hvec}(\mathbf{B})$ is an $R(R+1)/2$ dimensional vector.

To make (24) more concise, we stack all the $R(R-1)/2$ vectors $\text{svec}(\mathbf{R}_{st}^{ij})$, $1 \le s < t \le R$, into the $\frac{R(R+1)}{2} \times \frac{R(R-1)}{2}$ dimensional matrix

$$\mathbf{C}^{ij} = \left[\text{svec}\left(\mathbf{R}_{12}^{ij}\right), \ldots, \text{svec}\left(\mathbf{R}_{(R-1)R}^{ij}\right)\right]. \tag{25}$$

Furthermore, we stack all the M matrices $\mathbf{C}^{ij}$, $(i,j) \in \mathcal{I}$, into the $\frac{R(R+1)}{2} \times M\frac{R(R-1)}{2}$ matrix $[\mathbf{C}^{12}, \ldots, \mathbf{C}^{ij}, \ldots]$ and define the $\frac{R(R+1)}{2} \times \frac{R(R+1)}{2}$ dimensional matrix

$$\mathbf{C} = [\mathbf{C}^{12}, \ldots, \mathbf{C}^{ij}, \ldots][\mathbf{C}^{12}, \ldots, \mathbf{C}^{ij}, \ldots]^T. \tag{26}$$

Accordingly, (24) becomes $f(\mathbf{b}) = \mathbf{b}^T\mathbf{C}\mathbf{b}$ and the optimization (19) is changed to the vector form

$$\min \ f(\mathbf{b}) = \mathbf{b}^T\mathbf{C}\mathbf{b}, \quad \text{s.t.} \quad \text{rank}(\text{unhvec}(\mathbf{b})) = 1. \tag{27}$$

Lemma 2: If the columns of $\mathbf{C}^{ij}$ are linearly independent, then the null space of each matrix $\mathbf{C}^{ij}$, denoted as $\mathcal{N}\{\mathbf{C}^{ij}\}$, is spanned by the vectors $\mathbf{b}_k = \text{hvec}(\mathbf{w}_k^T \mathbf{w}_k)$, $k = 1, \ldots, R$, where $\mathbf{w}_k^T$ are source extraction vectors (and joint generalized eigenvectors).

Proof: If the columns of $\mathbf{C}^{ij}$ are linearly independent, then $\text{rank}(\mathbf{C}^{ij}) = R(R-1)/2$. Hence, the dimension of $\mathcal{N}\{\mathbf{C}^{ij}\}$ is R. According to (18) and Definition 2, we can see that $(\mathbf{C}^{ij})^T \mathbf{b}_k = \mathbf{0}$. Since the source extraction vectors $\mathbf{w}_k^T$ are linearly independent, so do the R vectors $\mathbf{b}_k$. Therefore, we conclude that $\mathcal{N}\{\mathbf{C}^{ij}\}$ is spanned by $\mathbf{b}_k$. ■

Lemma 2 shows that the optimal solutions to (27) are the joint generalized eigenvectors (source extraction vectors). In addition, Lemma 2 provides us with a guideline to select covariance matrices $\mathbf{R}_i$. One drawback of many eigen-decomposition-based source separation algorithms is the absence of such guidelines. Because each $\mathbf{C}_{ij}$ (and thus $\mathbf{C}$) has a null space with dimension R, the sum of the R least singular values of the matrix $\mathbf{C}$ should be zero. Therefore, in practical applications when the noise is not strictly white or the matrices $\mathbf{D}_i$ in (2) are not strictly diagonal, we can select the optimal set of covariance matrices $\mathbf{R}_i$ (via choosing time lags τ_i) by minimizing the sum of R least singular values of $\mathbf{C}$. The effectiveness of this guideline is verified by simulations in Section V-E.

B. Deflationary Algorithm for Sequential BSE and for the Estimation of Joint Generalized Eigenvectors

From (27), the gradient of $f(\mathbf{b})$ can be found as

$$\nabla f(\mathbf{b}) \stackrel{\triangle}{=} \frac{\partial f(\mathbf{b})}{\partial \mathbf{b}} = 2\mathbf{C}\mathbf{b}. \tag{28}$$

The updating rule is thus

$$\mathbf{b}^{(\ell)} = \mathbf{b}^{(\ell-1)} - 2\alpha^{(\ell-1)}\mathbf{C}\mathbf{b}^{(\ell-1)} \tag{29}$$

where the step size $\alpha^{(\ell-1)}$ can be optimized by exact line search

$$\begin{aligned}\alpha^{(\ell-1)} &= \arg\min_{\alpha>0} f(\mathbf{b}^{(\ell-1)} - 2\alpha\mathbf{C}\mathbf{b}^{(\ell-1)}) \\ &= \frac{(\mathbf{C}\mathbf{b}^{(\ell-1)})^T\mathbf{C}\mathbf{b}^{(\ell-1)}}{2(\mathbf{C}\mathbf{b}^{(\ell-1)})^T\mathbf{C}\mathbf{C}\mathbf{b}^{(\ell-1)}}.\end{aligned} \tag{30}$$

In each iteration, after obtaining $\mathbf{b}^{(\ell)}$, we conduct matrix shrinkage and normalization via

$$\mathbf{B}^{(\ell)} = \frac{S_v(\text{unhvec}(\mathbf{b}^{(\ell)}))}{\|S_v(\text{unhvec}(\mathbf{b}^{(\ell)}))\|_F}. \tag{31}$$

The normalization is necessary to avoid the trivial solution $\mathbf{B} = \mathbf{0}$. Upon convergence, we can calculate a joint generalized eigenvector (source extraction vector) from $\mathbf{B}^{(\ell)}$ as the eigenvector corresponding to the largest eigenvalue.

To estimate another joint generalized eigenvector, we apply deflation operation and run the iterations again. The overall procedure is outlined in Algorithm 1.

This algorithm estimates all the joint generalized eigenvectors (and source extraction vectors) sequentially, one in each stage. It runs R stages to estimate all the R vectors. For BSE applications, it can terminate early when the desired sources have been extracted. In each stage k, the algorithm runs iteratively the gradient search operation and the matrix shrinkage operation until the change in $\mathbf{B}^{(\ell)}$ is less than a pre-defined small threshold ζ.

Algorithm 1 : Estimate $\mathbf{w}_k^T$ and Extract Source $s_k(t)$ in Stage $k = 1, \ldots, R$.

Input: $\bar{\mathbf{x}}(t)$: observation signals; v, ζ: thresholds;
1: if $k = 1$: choose τ_i, find $\mathbf{R}_i$, $\mathbf{R}_{st}^{ij}$, $\mathbf{C}_1$ as per (5) (15) (26);
2: if $k > 1$: calculate $\mathbf{C}_k$ as per (33);
3: Initialize $\mathbf{B}^{(0)}$ and construct $\mathbf{b}^{(0)} = \text{hvec}(\mathbf{B}^{(0)})$;
4: $\ell \leftarrow 1$;
5: **repeat** (as per (29), (30), (31))
6: compute gradient: $\mathbf{g}^{(\ell-1)} = 2\mathbf{C}_k\mathbf{b}^{(\ell-1)}$;
7: compute step size: $\alpha^{(\ell-1)} = \frac{\mathbf{g}^{(\ell-1)T}\mathbf{g}^{(\ell-1)}}{2\mathbf{g}^{(\ell-1)T}\mathbf{C}_k\mathbf{g}^{(\ell-1)}}$;
8: update: $\mathbf{h}^{(\ell)} = \mathbf{b}^{(\ell-1)} - \alpha^{(\ell-1)}\mathbf{g}^{(\ell-1)}$;
9: unstack vector into matrix: $\mathbf{H}^{(\ell)} = \text{unhvec}(\mathbf{h}^{(\ell)})$
10: shrinkage: $\mathbf{B}^{(\ell)} = S_v(\mathbf{H}^{(\ell)})/\|S_v(\mathbf{H}^{(\ell)})\|_F$;
11: stack matrix into vector: $\mathbf{b}^{(\ell)} = \text{hvec}(\mathbf{B}^{(\ell)})$;
12: $\ell \leftarrow \ell + 1$;
13: **until** $\|\mathbf{B}^{(\ell)} - \mathbf{B}^{(\ell-1)}\|_F \le \zeta$;
14: estimate $\hat{\mathbf{w}}_k^T$ as leading eigenvector of $\mathbf{B}^{(\ell)}$;
15: calculate $\hat{s}_k(t) = \hat{\mathbf{w}}_k\bar{\mathbf{x}}(t)$.
Ouput: $\hat{\mathbf{w}}_k^T$ and $\hat{s}_k(t)$.

Deflation is applied to guarantee that the full set of orthogonal joint generalized eigenvectors are estimated. The deflation operation is conducted via determining the matrices $\mathbf{C}_k$ and $\mathbf{B}^{(0)}$. In stage $k = 1$, since no deflation is necessary, $\mathbf{C}_1 = \mathbf{C}$ and is calculated directly by (26). $\mathbf{B}^{(0)}$ is initialized as a random symmetric matrix with rank 1. In each of the subsequent stages $k \ge 2$, the targeting vector $\mathbf{w}_k^T$ should be orthogonal to all the vectors $\hat{\mathbf{w}}_1, \ldots, \hat{\mathbf{w}}_{k-1}$ estimated in previous stages, i.e.,

$$\hat{\mathbf{w}}_i\mathbf{w}_k^T = 0, \quad i = 1, \ldots, k-1. \tag{32}$$

This becomes $(\text{svec}(\hat{\mathbf{W}}_i))^T\mathbf{b} = 0$ in vector form, where $\mathbf{b} = \text{hvec}(\mathbf{w}_k^T\mathbf{w}_k)$ and $\hat{\mathbf{W}}_i = \hat{\mathbf{w}}_i^T\hat{\mathbf{w}}_i$. The idea of deflation is to look for a vector $\mathbf{b}$ that is orthogonal to $\mathbf{C}$ and all the estimated matrices $\text{svec}(\hat{\mathbf{W}}_i)(\text{svec}(\hat{\mathbf{W}}_i))^T$, $i = 1, \ldots, k-1$. This can be implemented by replacing the matrix $\mathbf{C}$ of (27) with

$$\mathbf{C}_k = \mathbf{C}_{k-1} + \text{svec}(\hat{\mathbf{W}}_{k-1})(\text{svec}(\hat{\mathbf{W}}_{k-1}))^T, \quad k \ge 2. \tag{33}$$

Furthermore, to speed up the gradient search, we can initialize $\mathbf{B}^{(0)}$ within the subspace orthogonal to $\hat{\mathbf{w}}_1, \ldots, \hat{\mathbf{w}}_{k-1}$. Considering the matrix $\hat{\mathbf{W}}^{(k-1)} = \left[\hat{\mathbf{w}}_1^T, \ \ldots, \ \hat{\mathbf{w}}_{k-1}^T\right]$, we conduct EVD

$$\hat{\mathbf{W}}^{(k-1)}(\hat{\mathbf{W}}^{(k-1)})^T = [\mathbf{U}_e, \ \mathbf{U}^{(k-1)}] \begin{bmatrix} \mathbf{\Sigma}_e & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} [\mathbf{U}_e, \ \mathbf{U}^{(k-1)}]^T. \tag{34}$$

Since $\mathbf{U}^{(k-1)} \in \mathbb{R}^{R\times(R-k+1)}$ is orthogonal to $\hat{\mathbf{w}}_1^T, \ldots, \hat{\mathbf{w}}_{k-1}^T$, the vector $\mathbf{w}_k^T$ is in the space spanned by the columns of $\mathbf{U}^{(k-1)}$. Therefore, we can initialize by $\mathbf{B}^{(0)} = \mathbf{U}^{(k-1)}\mathbf{c}\mathbf{c}^T\mathbf{U}^{(k-1)T}$ with a random vector $\mathbf{c} \in \mathbb{R}^{(R-k+1)\times 1}$.

The major computational complexity in each iteration comes from updating the $R(R+1)/2$ dimensional vector $\mathbf{b}$. The $MR(R-1)/2$ matrices $\mathbf{R}_{st}^{ij}$ are used just once at the beginning when calculating $\mathbf{C}_1$. All the subsequent iterations and stages involve only a single matrix $\mathbf{C}_k$. The complexity in each iteration is thus $O(R^4)$, dominated by the matrix-vector multiplication $\mathbf{C}_k\mathbf{b}^{(\ell-1)}$. This compares favorably with the common complexity $O(NR^3)$ of existing AJD methods [40]. This also compares favorably with the complexity $O(MR^2 \times R^3)$ of (21). Note that in many practical applications, the number of sources R is usually small while $N \gg R$ and $M \leq N(N-1)/2$ can be much larger, especially if we want to use a large number of covariance matrices to enhance estimation accuracy.

C. Convergence Analysis

Let us first analyze the operation of matrix shrinkage and normalization (31), which in Algorithm 1 is described as $\mathbf{B}^{(\ell)} = S_v(\mathbf{H}^{(\ell)})/\|S_v(\mathbf{H}^{(\ell)})\|_F$. The objective of matrix shrinkage is to make the matrix $\mathbf{B}^{(\ell)}$ closer to rank 1 than $\mathbf{H}^{(\ell)}$. Let the singular values of $\mathbf{H}^{(\ell)}$ and $\mathbf{B}^{(\ell)}$ be $\sigma_{h,1} \geq \cdots \geq \sigma_{h,R}$ and $\sigma_{b,1} \geq \cdots \geq \sigma_{b,R}$, respectively. Their distances from rank 1 can be measured by $\mu_h = \sigma_{h,1}/\sum_{i=2}^{R}\sigma_{h,i}$ and $\mu_b = \sigma_{b,1}/\sum_{i=2}^{R}\sigma_{b,i}$, respectively. Larger value means closer to rank 1. We define $\sigma_{h,R+1} = 0$ for notational convenience.

Lemma 3: The operation of matrix shrinkage and normalization makes $\mathbf{B}^{(\ell)}$ closer to rank 1 than $\mathbf{H}^{(\ell)}$, i.e., $\mu_b > \mu_h$, if the threshold v satisfies

$$\sigma_{h,n} \geq v \geq \sigma_{h,n+1},$$
$$v < \frac{\sum_{i=n+1}^{R}\sigma_{h,i}}{\left(\sigma_{h,1}^{-1}(\sum_{i=2}^{R}\sigma_{h,i}) - n + 1\right)^{+}}, \tag{35}$$

for an integer $n = 1, \ldots, R$.

Proof: See Appendix D. ■

A desirable property is that the right-hand-side of (35), i.e., the upper bound of v, increases when $\mathbf{H}^{(\ell)}$ is closer to rank 1. This means that we just need to calculate this upper bound once using $\mathbf{H}^{(1)}$ obtained in the first iteration. As long as we choose v under this upper bound, Lemma 3 will guarantee that the operation of matrix shrinkage and normalization makes $\mathbf{B}^{(\ell)}$ more and more close to rank 1.

Theorem 3: With a relatively small v that satisfies (35) and a step size α that satisfies $0 < \alpha < \frac{1}{\lambda_{\max}}$, where $\lambda_{\max}$ is the maximum eigenvalue of $\mathbf{C}_k$, the Algorithm 1 converges to $\mathbf{B}^*$.

Proof: See Appendix E. ■

From $\lambda_{\max} = \max_{\mathbf{z}}(\mathbf{z}^T\mathbf{C}_k\mathbf{z})/(\mathbf{z}^T\mathbf{z})$, it can be readily shown that $\mathbf{b}^T\mathbf{C}^T\mathbf{C}\mathbf{C}\mathbf{b}/(\mathbf{b}^T\mathbf{C}^T\mathbf{C}\mathbf{b}) > \lambda_{\max}/2$ for random $\mathbf{b}$ and thus the optimized $\alpha^{(\ell-1)}$ in (30) always satisfies $0 < \alpha^{(\ell-1)} < 1/\lambda_{\max}$.

V. Simulations

In this section, we evaluate the performance of our proposed BSE algorithm by simulations. We compare it with six typical BSS or BSE algorithms, i.e., SOBI [29], FastICA [67], JADE [31], WASOBI [68], BGSEP [69], and RGD (relative gradient decent) [34]. For WASOBI and BGSEP, we used the MATLAB code obtained from [70]. For FastICA, we used the MATLAB code obtained 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 source code of our proposed BSE algorithm can be obtained at `http://www.ws.binghamton.edu/li/ BSE-Zhang-Li-Peng2017.zip`. Note that the SOBI, JADE, WASOBI, BGSEP and RGD algorithms all use joint diagonalization techniques to recover all the source signals simultaneously.

As performance measure, we used the performance index

$$\text{PI} = \frac{1}{R(R-1)}\sum_{i=1}^{R}\left\{\left(\sum_{k=1}^{R}\frac{|\hat{g}_{ik}|}{\max_j|\hat{g}_{ij}|} - 1\right) + \left(\sum_{k=1}^{R}\frac{|\hat{g}_{ki}|}{\max_j|\hat{g}_{ji}|} - 1\right)\right\} \tag{36}$$

where $\hat{g}_{ij}$ is the (i,j)-element of the mixing-separating matrix $\hat{\mathbf{G}} = \hat{\mathbf{W}}\mathbf{U}\mathbf{A}$. This index measures to what extent $\hat{\mathbf{G}}$ is close to a generalized permutation matrix [1]. The smaller the value of PI, the better the separation performance.

The mixing matrix $\mathbf{A}$ and the noise were randomly generated in each run of simulations. Our new algorithm was initialized with $\mathbf{B}^{(0)} = \mathbf{1}_R\mathbf{1}_R^T$ in stage $k = 1$, where $\mathbf{1}_R$ is an R dimensional vector whose elements are all one. In all other stages $k \geq 2$, our algorithm was initialized with $\mathbf{c} = \mathbf{1}_{R-k+1}$.

A. Extraction of AR Signals

In this experiment, each source signal was modeled as an Lth order autoregressive (AR) process driven by the zero-mean Gaussian white noise with unit variance, i.e.,

$$s_i(t) = \sum_{l=1}^{L}h_i(l)s_i(t-l) + w_i(t), \quad 1 \leq i \leq R \tag{37}$$

where $h_i(l)$ are AR coefficients. The values of $w_i(t)$ were generated independently from the standard normal distribution. For fair comparison with [68], we adopted its model parameters where $h_i(l), l = 1, \ldots, L-1$, were fixed to 0, while $h_i(L)$ was randomly generated in the interval $(-1, 1)$. As pointed out in [68], $h_i(L)$ can be used to measure whether source signals are easy to separate or not. When $h_i(L)$ is close to zero, it is difficult to separate the mixed source signals because they have similar correlations or power spectra. On the other hand, when $|h_i(L)|$ is close to one, sources become relatively easy to separate thanks to their distinct correlations or power spectra.

We conducted simulations with parameters $J = 5$, $R = 3$, $N = 10$, and $L = 5$. We used time lags $\tau_i = i$, $i = 1, \ldots, 10$, to generate covariance matrices $\mathbf{R}_i$. We also added white noise so as to create mixtures $\mathbf{x}(t)$ with various signal-to-noise ratios (SNRs), where SNR was defined as $\text{SNR} = 10\log_{10}\frac{E\{\|\mathbf{A}\mathbf{s}(t)\|^2\}}{E\{\|\mathbf{n}(t)\|^2\}}$. We used 100 independent runs to calculate the average PI. Simulation results are shown in Fig. 2.

It can be seen that WASOBI had the best performance when SNR was extremely high. However, both our proposed algorithm

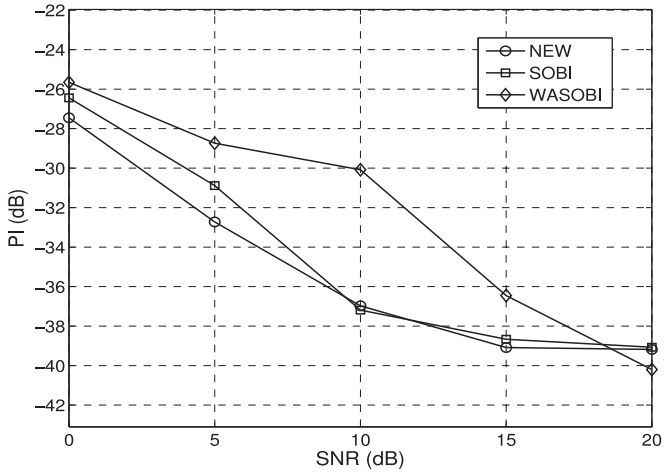

Fig. 2. Performance index (PI) as function of SNR in the experiment of extracting AR signals. Our proposed algorithm (NEW) outperformed the other two comparing algorithms when SNR was not extremely high.

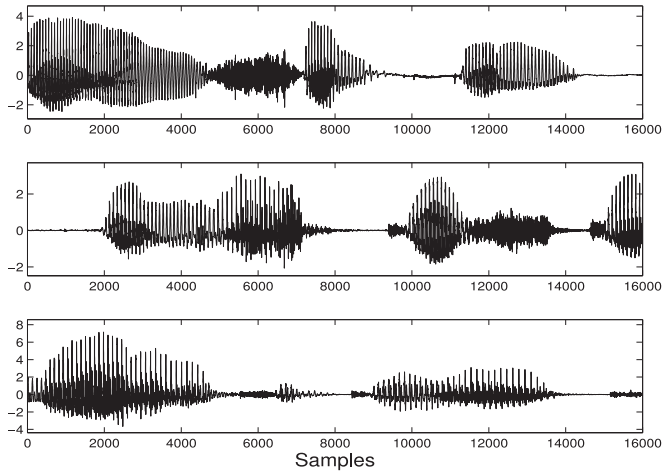

Fig. 3. Three speech source signals.

and SOBI had better performance than WASOBI when SNR was not extremely high. This showed that WASOBI was sensitive to noise. In addition, our proposed algorithm was slightly better than SOBI.

We also measured the runtime of these algorithms on a 2.40-GHz PC with 4 GB RAM. The average runtimes were 0.020725, 0.011602 and 0.023464 seconds for our proposed algorithm, SOBI, and WASOBI, respectively. Note that SOBI and WASOBI estimated all the source signals simultaneously while our proposed algorithm extracted source signals one by one. Hence, when the aim is just to extract a subset of the source signals, the runtime of our proposed algorithm will be much more competitive.

B. Extraction of Speech Signals

In this experiment, we considered $R = 3$ speech source signals. Each of them was 1 second long (see Fig. 3), and was obtained from [71] with truncation. Speech signals are known to be non-Gaussian, non-white, and non-stationary. For the algorithms compared in this experiment, JADE and FastICA exploit

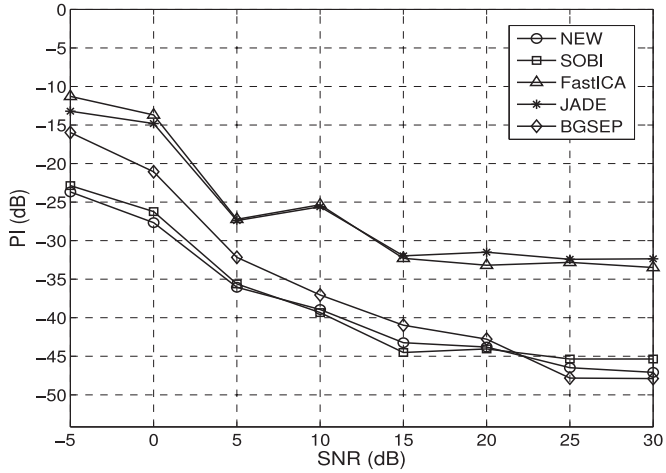

Fig. 4. Performance index (PI) as function of SNR in the experiment of speech signal extraction. Our proposed algorithm (NEW) had better performance and was more robust in low SNRs.

the non-Gaussianity of the source signals while our proposed algorithm and SOBI utilize the non-whiteness of the source signals. The BGSEP [69], [72] exploits the non-stationarity of the source signals instead.

The number of observation signals was $J = 5$. The number of covariance matrices was $N = 10$. For BGSEP, the number of blocks of signals was set to 10. For fair comparison, our proposed algorithm and SOBI were simulated with the same set of covariance matrices. We used 100 independent runs to calculate the average PI. Simulation results are shown in Fig. 4.

As it can be seen, the performance of FastICA and JADE was much worse than the other algorithms. When the SNR was not extremely high, the performance of BGSEP was worse than our algorithm and SOBI. In addition, our proposed algorithm was slightly better than SOBI, especially at low SNR, e.g., SNR $\in [-5,\ 10]$ dB. This means that our proposed algorithm is more robust to heavy noise than the other algorithms.

To explain why the BGSEP is not as robust to noise as our proposed algorithm, let us denote by $\mathbf{x}_k(t)$, $\mathbf{s}_k(t)$ and $\mathbf{n}_k(t)$ the kth block of the observation signal, source signal and noise, respectively. Define the covariance matrix of $\mathbf{x}_k(t)$ with zero time-lag as

$$\bar{\mathbf{C}}_{x,k} = \mathbf{A}\mathbf{D}_k\mathbf{A}^T + \mathbf{R}_{nn,k} \tag{38}$$

where $\mathbf{D}_k \stackrel{\triangle}{=} E\{\mathbf{s}_k(t)\mathbf{s}_k{}^T(t)\}$ and $\mathbf{R}_{nn,k} \stackrel{\triangle}{=} E\{\mathbf{n}_k(t)\mathbf{n}_k^T(t)\}$. The BGSEP algorithm uses the zero time-lagged covariance matrices $\bar{\mathbf{C}}_{x,1}, \ldots, \bar{\mathbf{C}}_{x,M}$ to recover the source signals. As we can see, such covariance matrices cannot avoid the influence of noise. In contrast, our proposed algorithm uses non-zero time-lagged covariance matrices $\mathbf{C}_{x,i}$ which are not susceptible to white noise theoretically. This is one of the primary reasons why our proposed algorithm outperformed BGSEP when the SNR was not extremely high.

C. Robustness to Ill-Conditioned Mixing Matrices

For the mixing matrix $\mathbf{A}$, its condition number can be defined as $\kappa(\mathbf{A}) = \frac{\sigma_1}{\sigma_R}$, where σ_1 and σ_R are the largest and the smallest

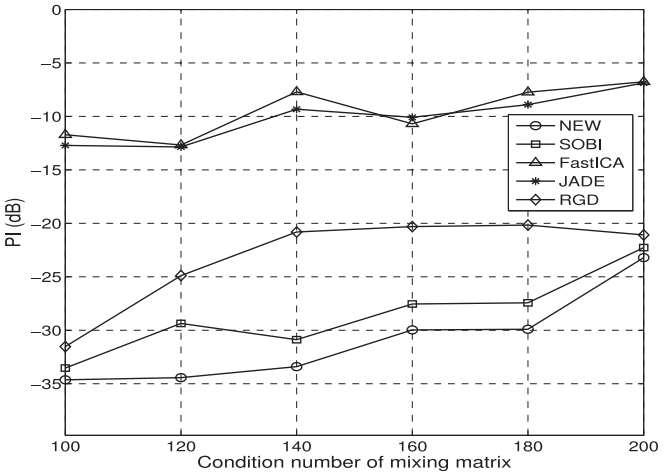

Fig. 5. Performance index (PI) as function of condition number of the mixing matrix $\mathbf{A}$. Our proposed algorithm (NEW) was more robust and outperformed the other algorithms.

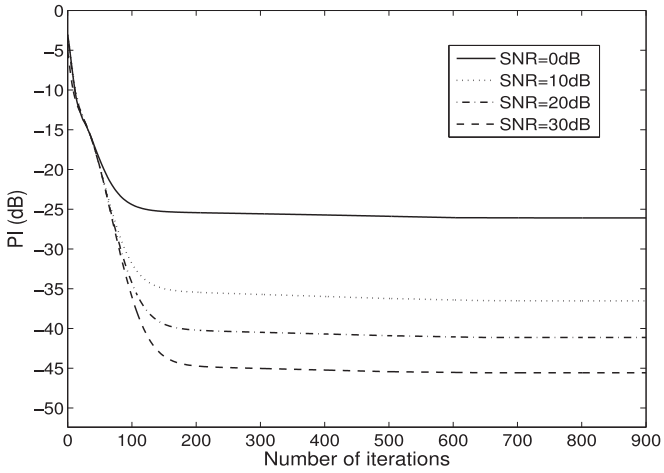

Fig. 6. Performance index (PI) as function of number of iterations under various SNRs. $\alpha = 0.05, v = 0.0005$.

TABLE I
LARGEST EIGENVALUES $\lambda_{\max}$ OF $\mathbf{C}_k$ UNDER VARIOUS SNRS

SNR (dB)	0	5	10	15	20	25	30
λ_{max}	6.44	6.42	6.19	6.16	6.13	6.10	6.09

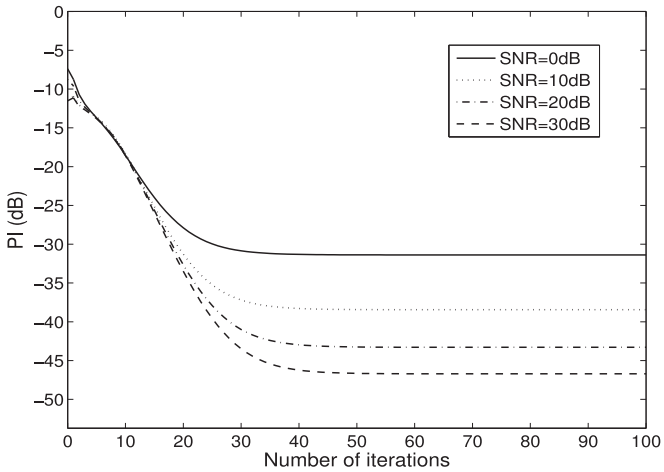

Fig. 7. Performance index (PI) as function of number of iterations under various SNRs. $\alpha = 0.1, v = \max(v_0(\gamma_v)^{\ell-1}, 0.00001)$.

singular values of $\mathbf{A}$, respectively [73]. Large $\kappa(\mathbf{A})$ means that $\mathbf{A}$ is near singular and thus ill-conditioned.

In order to evaluate the robustness of our algorithm to ill-conditioned mixing matrices, in this experiment, we generated $\mathbf{A}$ with $\kappa(\mathbf{A})$ varying from 100 to 200. With such large $\kappa(\mathbf{A})$ values, the source extraction problem was highly ill-conditioned.

The source signals used in this experiment were the same speech signals as those in Section V-B. The simulation results shown in Fig. 5 demonstrated that our proposed algorithm was more robust and outperformed the other algorithms. In this ill-conditioned situation, it was hard for FastICA and JADE to separate source signals. In contrast, our proposed algorithm, SOBI, and RGD still had good performance with PI ranging from -20 dB to -35 dB. All of them were robust to the ill-conditioned mixing matrices, and our proposed algorithm had the best performance.

D. Evaluation of Convergence

In this experiment, we set the number of observation signals to $J = 6$ and the number of covariance matrices to $N = 10$. We used again the three speech source signals studied in Section V-B. Table I lists the largest eigenvalues $\lambda_{\max}$ of the matrices $\mathbf{C}_k$ under various SNRs.

According to the convergence results given in Theorem 3, we require $0 < \alpha < \frac{1}{\lambda_{\max}}$ in order to guarantee the convergence of the proposed algorithm. First, we adopted the constant step size $\alpha = 0.05$, which satisfied the convergence condition. The threshold of matrix shrinkage $S_v(\cdot)$ was chosen as $v = 0.0005$. Simulation results shown in Fig. 6 demonstrated that the proposed algorithm had the guaranteed convergence.

Next, to evaluate the connection between the parameters α, v and the convergence rate, we set the step size to $\alpha = 0.1$ and $v = \max(v_0(\gamma_v)^{\ell-1}, 0.00001)$, where $v_0 = 0.05$ and $\gamma_v = 0.995$. Note that this large α still satisfied the convergent condition $0 < \alpha < 1/\lambda_{\max}$. The parameter v gradually decreased from a big value 0.05 to a very small constant 0.00001. Simulation results in Fig. 7 indicated that larger step size α and varying threshold v could substantially increase the convergence speed and enhance the source extraction accuracy.

E. Covariance Matrix Optimization

Many BSS or BSE algorithms use the covariance matrices calculated simply with time lags $\tau_i \in \{1, \ldots, N\}$. We denote by T_1 this set of covariance matrices. In this experiment, we evaluated the performance of our algorithm when using a set of covariance matrices obtained with a slightly optimized set of time lags. The

TABLE II
PERFORMANCE INDEX WHEN USING COVARIANCE MATRICES WITHOUT OPTIMIZATION (T_1) OR WITH OPTIMIZATION (T_2)

SNR(dB)	−5	0	5	10	15	20	25	30
T_1:PI(dB)	–21	–29	–29	–39	–47	–47	–50	–48
T_2:PI(dB)	–25	–31	–36	–40	–49	–48	–50	–50

criterion for time lag optimization was to minimize the sum of R least singular values of $\mathbf{C}$. We denote by T_2 this set of covariance matrices. Note that T_2 may vary in different runs of simulations due to the variety of noise. An example of the optimized time lags was $\tau_i \in \{1, 2, 3, 9, 10, 11, 12, 13, 18, 20\}$. The parameters J, R, N and the source signals used in this experiment were the same as those in the previous subsection. From the simulation results listed in Table II, it can be seen that such an optimization could enhance the performance by several decibels.

VI. CONCLUSION

In this paper we developed a new sequential blind source extraction (BSE) approach based on the joint generalized eigenvectors of a series of covariance matrix pencils. We proved that the joint generalized eigenvectors are the same as source extraction vectors, and can be estimated from an optimization involving a quadratic cost function and a unit-rank matrix constraint. We used gradient search and matrix shrinkage to solve this optimization, and implemented it as an efficient algorithm based on the symmetry-preserving vectorization and deflation techniques. The merits of the proposed algorithm include superior performance, robustness to noise, and robustness to ill-conditioned mixing matrices. Simulations were conducted to demonstrate that this algorithm compared favorably to some typical BSS and BSE algorithms.

APPENDIX A
PROOF OF THEOREM 1

From (9) and (10), because $\mathbf{w}_k^T$ satisfies (10) for all $(i, j) \in \mathcal{I}$, each source extraction vector $\mathbf{w}_k^T$ is also a joint generalized eigenvector. Since the dimension of the generalized eigenvector space for each matrix pencil is R only, there is no other join generalized eigenvector that is independent from $\mathbf{w}_k^T$. Therefore, we just need to show that linear combinations of $\mathbf{w}_k^T$ cannot be joint generalized eigenvectors.

Sufficiency: To show the sufficiency of Theorem 1, we will prove that if (11) is satisfied, then none of the linear combinations of $\mathbf{w}_k^T$ with at least two nonzero combining coefficients can be a joint generalized eigenvector. We do it by contradiction.

Assume the linear combination $\mathbf{v}^T = \sum_{k=1}^R \alpha_k \mathbf{w}_k^T$ be a joint generalized eigenvector instead. Without any loss of generality, we assume $\alpha_1 \neq 0$ and $\alpha_\ell \neq 0$ for some $2 \leq \ell \leq R$. According to (11), for any index m, $2 \leq m \leq R$, there exists an index pair (i, j) such that

$$d_{im} d_{j1} - d_{i1} d_{jm} \neq 0. \tag{39}$$

For this index pair (i, j), taking into account (9), we have

$$\mathbf{R}_i \mathbf{v}^T = \mathbf{R}_i \sum_{k=1}^R \alpha_k \mathbf{w}_k^T = \sum_{k=1}^R \alpha_k d_{ik} \gamma_k \mathbf{U}\mathbf{a}_k \tag{40}$$

and

$$\mathbf{R}_j \mathbf{v}^T = \mathbf{R}_j \sum_{k=1}^R \alpha_k \mathbf{w}_k^T = \sum_{k=1}^R \alpha_k d_{jk} \gamma_k \mathbf{U}\mathbf{a}_k. \tag{41}$$

Let λ be the generalized eigenvalue of the matrix pencil $(\mathbf{R}_i, \mathbf{R}_j)$ corresponding to the generalized eigenvector $\mathbf{v}^T$. We have

$$\mathbf{R}_i \sum_{k=1}^R \alpha_k \mathbf{w}_k^T = \lambda \mathbf{R}_j \sum_{k=1}^R \alpha_k \mathbf{w}_k^T. \tag{42}$$

Combining (40)–(42), we have

$$\sum_{k=1}^R \alpha_k (d_{ik} - \lambda d_{jk}) \gamma_k \mathbf{U}\mathbf{a}_k = \mathbf{0} \tag{43}$$

where $\mathbf{0}$ denotes zero vector. Because $\mathbf{UA}$ is a square full column rank matrix, its column vectors $\mathbf{U}\mathbf{a}_k, k = 1, \ldots, R$, are linearly independent. Hence, we have

$$\alpha_k (d_{ik} - \lambda d_{jk}) \gamma_k = 0, \quad 1 \leq k \leq R. \tag{44}$$

Taking into account $\gamma_k \neq 0$, we have

$$\alpha_k (d_{ik} - \lambda d_{jk}) = 0, \quad 1 \leq k \leq R. \tag{45}$$

Since $\alpha_1 \neq 0$, this leads to $d_{i1} - \lambda d_{j1} = 0$ which is

$$d_{i1} = \lambda d_{j1}. \tag{46}$$

With (46), we can change (39) into $(d_{im} - \lambda d_{jm}) d_{j1} \neq 0$. Therefore, we have $d_{im} - \lambda d_{jm} \neq 0$ for all $2 \leq m \leq R$. Then from (45), we must have $\alpha_k = 0$ for all $2 \leq k \leq R$, which is a contradiction to the assumption $\alpha_\ell \neq 0$. This proves the sufficiency.

Necessity: The necessity can be established by proving that if there exists an index pair m and n, $1 \leq m \neq n \leq R$, that satisfy $d_{im} d_{jn} - d_{in} d_{jm} = 0$ for all $(i, j) \in \mathcal{I}$, then we can find a linear combination (with at least two nonzero coefficients) of the vectors $\mathbf{w}_k^T$ that is a joint generalized eigenvector.

Since the source extraction vectors $\mathbf{w}_m^T$ and $\mathbf{w}_n^T$ are joint generalized eigenvectors, according to (9), for all $(i, j) \in \mathcal{I}$ we have

$$\mathbf{R}_i \mathbf{w}_m^T = d_{im} \gamma_m \mathbf{U}\mathbf{a}_m, \quad \mathbf{R}_j \mathbf{w}_m^T = d_{jm} \gamma_m \mathbf{U}\mathbf{a}_m, \tag{47}$$

$$\mathbf{R}_i \mathbf{w}_n^T = d_{in} \gamma_n \mathbf{U}\mathbf{a}_n, \quad \mathbf{R}_j \mathbf{w}_n^T = d_{jn} \gamma_n \mathbf{U}\mathbf{a}_n. \tag{48}$$

Therefore, for the linear combination $\mathbf{v}^T = \alpha \mathbf{w}_m^T + \beta \mathbf{w}_n^T$, where $\alpha \neq 0$, $\beta \neq 0$, we have

$$\mathbf{R}_i \mathbf{v}^T = \alpha d_{im} \gamma_m \mathbf{U}\mathbf{a}_m + \beta d_{in} \gamma_n \mathbf{U}\mathbf{a}_n, \tag{49}$$

$$\mathbf{R}_j \mathbf{v}^T = \alpha d_{jm} \gamma_m \mathbf{U}\mathbf{a}_m + \beta d_{jn} \gamma_n \mathbf{U}\mathbf{a}_n. \tag{50}$$

Next, we exploit the condition $d_{im} d_{jn} = d_{in} d_{jm}$ for analysis.

Case 1: $d_{in} = d_{jn} = 0$. In this case we have

$$\mathbf{R}_i \mathbf{v}^T = \alpha d_{im} \gamma_m \mathbf{U}\mathbf{a}_m, \quad \mathbf{R}_j \mathbf{v}^T = \alpha d_{jm} \gamma_m \mathbf{U}\mathbf{a}_m. \tag{51}$$

Because (51) is valid for all $(i,j) \in \mathcal{I}$, the linear combination $\mathbf{v}^T$ is a joint generalized eigenvector.

Case 2: $d_{in} \neq 0$ and $d_{jn} \neq 0$. In this case we can change (49) and (50) into

$$\begin{aligned}\mathbf{R}_i\mathbf{v}^T &= d_{in}\left(\alpha\frac{d_{im}}{d_{in}}\gamma_m\mathbf{U}\mathbf{a}_m + \beta\gamma_n\mathbf{U}\mathbf{a}_n\right)\\ &= d_{in}\left(\alpha\frac{d_{jm}}{d_{jn}}\gamma_m\mathbf{U}\mathbf{a}_m + \beta\gamma_n\mathbf{U}\mathbf{a}_n\right), \qquad (52)\\ \mathbf{R}_j\mathbf{v}^T &= d_{jn}\left(\alpha\frac{d_{jm}}{d_{jn}}\gamma_m\mathbf{U}\mathbf{a}_m + \beta\gamma_n\mathbf{U}\mathbf{a}_n\right). \qquad (53)\end{aligned}$$

This also implies that the linear combination $\mathbf{v}$ is a joint generalized eigenvector.

Case 3: Only one of d_{in} and d_{jn} is zero. Without any loss of generality, consider $d_{in} = 0$ and $d_{jn} \neq 0$. Then we have $d_{im} = 0$. From (49) and (50) we know that $\mathbf{v}^T$ is a joint generalized eigenvector with $\lambda_{i,j} = 0$.

Combining the above three cases, the necessity is proved.

Appendix B
Proof of Lemma 1

First, if $\mathbf{u}_i$ and $\mathbf{u}_j$ are proportional, i.e., $\mathbf{u}_i = \gamma\mathbf{u}_j$ for some scalar $\gamma \neq 0$, then $u_{is}u_{jt} - u_{it}u_{js} = \gamma u_{js}u_{jt} - \gamma u_{jt}u_{js} = 0$. Next, if $u_{is}u_{jt} - u_{it}u_{js} = 0$ for all $1 \leq s < t \leq R$, we need to prove that there exists a nonzero scalar γ such that $\mathbf{u}_i = \gamma\mathbf{u}_j$. Let $u_{ik} \neq 0$ for some k. Then we have $u_{jt} = \frac{u_{jk}}{u_{ik}}u_{it}$ for all $t > k$ and $u_{js} = \frac{u_{jk}}{u_{ik}}u_{is}$ for all $s < k$. Let $\gamma = \frac{u_{jk}}{u_{ik}}$. This gives $\mathbf{u}_i = \gamma\mathbf{u}_j$.

What remains is to show $\gamma \neq 0$, which means we must have $u_{jk} \neq 0$. This can be readily shown by contradiction. Specifically, assume $u_{jk} = 0$ instead. From $u_{ik}u_{jt} = u_{it}u_{jk}$ we see that $u_{jt} = 0$ for all $t > k$. In addition, from $u_{is}u_{jk} = u_{ik}u_{js}$ we find $u_{js} = 0$ for all $s < k$. This means that $\mathbf{u}_j = \mathbf{0}$, a contradiction.

Appendix C
Proof of Theorem 2

First, if $\mathbf{B} = \mathbf{w}_k^T\mathbf{w}_k$ where $\mathbf{w}_k^T$ is a joint generalized eigenvector (and is also a source extraction vector according to Theorem 1), then $\text{rank}(\mathbf{B}) = 1$ and $f(\mathbf{B}) = 0$. This also means that the minimum value of $f(\mathbf{B})$ is zero and is achievable.

Second, assume that by optimizing (19) we have found a symmetric matrix $\mathbf{B}$ that satisfies both $\text{rank}(\mathbf{B}) = 1$ and $f(\mathbf{B}) = 0$. It is easy to see that $\mathbf{B}$ satisfies (18). In addition, we can find a vector $\mathbf{w}^T$ such that $\mathbf{B} = \mathbf{w}^T\mathbf{w}$. We can see that $\mathbf{w}^T$ satisfies (17) and also (16). It can be readily verified that (16) is equivalent to (14) for real signals. Therefore, based on Lemma 1, $\mathbf{w}^T$ satisfies (10) and is a joint generalized eigenvector. Theorem 1 guarantees that $\mathbf{w}^T$ is also a source extraction vector.

Appendix D
Proof of Lemma 3

Obviously, normalization does not change the measure μ_b. Therefore, we just need to analyze matrix shrinkage. To make $\mu_b > \mu_h$, we require

$$\frac{(\sigma_{h,1} - v)^+}{\sum_{i=2}^R(\sigma_{h,i} - v)^+} > \frac{\sigma_{h,1}}{\sum_{i=2}^R\sigma_{h,i}}. \qquad (54)$$

Consider the case that the threshold v satisfies

$$\sigma_{h,n} \geq v \geq \sigma_{h,n+1} \qquad (55)$$

for some integer $n = 1, \ldots, R$. Note that $n = R$ means that v is no larger than any singular value. Note also that v should be smaller than $\sigma_{h,1}$ because otherwise the result of matrix shrinkage is zero according to (22)(23). With (55), we can change (54) to

$$\frac{\sigma_{h,1} - v}{\sum_{i=2}^n(\sigma_{h,i} - v)} > \frac{\sigma_{h,1}}{\sum_{i=2}^R\sigma_{h,i}}. \qquad (56)$$

Re-arranging the items in (56) we come at

$$\left(\frac{1}{\sigma_{h,1}}\sum_{i=2}^R\sigma_{h,i} - n + 1\right)v < \sum_{i=n+1}^R\sigma_{h,i}. \qquad (57)$$

If $(\frac{1}{\sigma_{h,1}}\sum_{i=2}^R\sigma_{h,i}) - n + 1 \leq 0$, then (57) holds for any v. If $(\frac{1}{\sigma_{h,1}}\sum_{i=2}^R\sigma_{h,i}) - n + 1 > 0$, then we have

$$v < \frac{\sum_{i=n+1}^R\sigma_{h,i}}{\sigma_{h,1}^{-1}\sum_{i=2}^R\sigma_{h,i} - n + 1}. \qquad (58)$$

Combining these two cases leads to (35).

Appendix E
Proof of Theorem 3

Without any loss of generality, assume a constant step size α for simplification. From (29), in stage k, we have

$$\mathbf{b}^{(\ell)} = (1 - 2\alpha\mathbf{C}_k)\bar{\mathbf{b}}^{(\ell-1)}, \qquad (59)$$

where

$$\bar{\mathbf{b}}^{(\ell-1)} = \text{hvec}(\bar{\mathbf{B}}^{(\ell-1)}) = \text{hvec}\left(\frac{S_v(\text{unhvec}(\mathbf{b}^{(\ell-1)}))}{\|S_v(\text{unhvec}(\mathbf{b}^{(\ell-1)}))\|_F}\right). \qquad (60)$$

Note that if the matrix shrinkage and normalization operation is not applied, $\bar{\mathbf{b}}^{(\ell-1)}$ should be replaced with $\mathbf{b}^{(\ell-1)}$, which gives

$$\mathbf{b}^{(\ell)} = (1 - 2\alpha\mathbf{C}_k)\mathbf{b}^{(\ell-1)}. \qquad (61)$$

With the EVD $\mathbf{C}_k = \mathbf{U}_c\text{diag}\{\lambda_1, \ldots, \lambda_{R(R+1)/2}\}\mathbf{U}_c^T$, we can rewrite it as

$$\mathbf{U}_c^T\mathbf{b}^{(\ell)} = \text{diag}\{1 - 2\alpha\lambda_1, \ldots, 1 - 2\alpha\lambda_{R(R+1)/2}\}\mathbf{U}_c^T\mathbf{b}^{(\ell-1)}. \qquad (62)$$

It is easy to see that the iteration (61) or (62) converges when $|1 - 2\alpha\lambda_i| < 1$, which leads to $0 < \alpha < 1/\lambda_{\max}$ [58]. However, (61) may possibly converge to the trivial zero solution instead of $\mathbf{B}^*$, which is why the matrix shrinkage and normalization operation is needed.

In the following, we will first show that the updating will lead $\mathbf{B}^{(\ell-1)}$ to the neighborhood of an optimal solution $\mathbf{B}^*$. Then, we will show that the algorithm can converge to $\mathbf{B}^*$.

First, since $\mathbf{b}^* = \text{hvec}(\mathbf{B}^*)$ and $\mathbf{C}_k\mathbf{b}^* = \mathbf{0}$, from (59) we can get

$$\begin{aligned}\mathbf{b}^{(\ell-1)} - \mathbf{b}^* &= (1-2\alpha\mathbf{C}_k)\bar{\mathbf{b}}^{(\ell-2)} - \mathbf{b}^* \\ &= (1-2\alpha\mathbf{C}_k)(\bar{\mathbf{b}}^{(\ell-2)} - \mathbf{b}^*) \\ &= (1-2\alpha\mathbf{C}_k)(\bar{\mathbf{b}}^{(\ell-2)} - \mathbf{b}^{(\ell-2)}) \\ &\quad + (1-2\alpha\mathbf{C}_k)(\mathbf{b}^{(\ell-2)} - \mathbf{b}^*). \end{aligned} \tag{63}$$

In (63), we can replace $\mathbf{b}^{(\ell-2)} - \mathbf{b}^*$ by a similar equation containing the item $\mathbf{b}^{(\ell-3)} - \mathbf{b}^*$, i.e., $(1-2\alpha\mathbf{C}_k)(\bar{\mathbf{b}}^{(\ell-3)} - \mathbf{b}^{(\ell-3)}) + (1-2\alpha\mathbf{C}_k)(\mathbf{b}^{(\ell-3)} - \mathbf{b}^*)$. Doing this operation iteratively we can finally describe $\mathbf{b}^{(\ell-1)} - \mathbf{b}^*$ by $\mathbf{b}^{(0)} - \mathbf{b}^*$. As a result, with some straightforward deductions, we can get

$$\begin{aligned}\mathbf{b}^{(\ell-1)} - \mathbf{b}^* &= \sum_{i=1}^{\ell-1}(1-2\alpha\mathbf{C}_k)^i(\bar{\mathbf{b}}^{(\ell-1-i)} - \mathbf{b}^{(\ell-1-i)}) \\ &\quad + (1-2\alpha\mathbf{C}_k)^{\ell-1}(\mathbf{b}^{(0)} - \mathbf{b}^*). \end{aligned} \tag{64}$$

According to the matrix shrinkage and normalization procedure, for the unit-norm matrices $\mathbf{B}^{(\ell-1)}$, we have $\|\bar{\mathbf{B}}^{(\ell-1)} - \mathbf{B}^{(\ell-1)}\| < Cv$ for some constant C. This means that

$$\begin{aligned}\|\mathbf{b}^{(\ell-1)} - \mathbf{b}^*\| &< \sum_{i=1}^{\ell-1}\|1-2\alpha\mathbf{C}_k\|^i Cv \\ &\quad + \|1-2\alpha\mathbf{C}_k\|^{\ell-1}\|\mathbf{b}^{(0)} - \mathbf{b}^*\| \end{aligned} \tag{65}$$

Therefore, if $0 < \alpha < 1/\lambda_{\max}$ and v is small, we can see that $\mathbf{b}^{(\ell-1)}$ is close to $\mathbf{b}^*$ for large enough ℓ. Note that according to (35), v can be very small. As a result, $\mathbf{B}^{(\ell-1)}$ falls within the neighborhood of $\mathbf{B}^*$. Since $\mathbf{B}^*$ has unit rank, the first (dominant) singular vector of $\mathbf{B}^{(\ell-1)}$ is close to that of $\mathbf{B}^*$.

Next, considering that

$$\mathbf{b}^{(\ell)} - \mathbf{b}^* = (1-2\alpha\mathbf{C}_k)(\bar{\mathbf{b}}^{(\ell-1)} - \mathbf{b}^*), \tag{66}$$

we have

$$\|\mathbf{b}^{(\ell)} - \mathbf{b}^*\| \le \|1-2\alpha\mathbf{C}_k\|\|\bar{\mathbf{b}}^{(\ell-1)} - \mathbf{b}^*\|. \tag{67}$$

To analyze $\|\bar{\mathbf{b}}^{(\ell-1)} - \mathbf{b}^*\|$, we consider the equivalent matrix form. According to Lemma 3, if v satisfies (35) and the dominant singular vector of $\mathbf{B}^{(\ell-1)}$ is close to that of $\mathbf{B}^*$, then

$$\left\| \frac{S_v(\mathbf{B}^{(\ell-1)})}{\|S_v(\mathbf{B}^{(\ell-1)})\|_F} - \mathbf{B}^* \right\| \le \|\mathbf{B}^{(\ell-1)} - \mathbf{B}^*\|. \tag{68}$$

This is because the matrix shrinkage operation makes the matrix $\mathbf{B}^{(\ell-1)}$ closer to the rank 1 matrix $\mathbf{B}^*$. With (68), we can change (67) into

$$\|\mathbf{b}^{(\ell)} - \mathbf{b}^*\| \le \|1-2\alpha\mathbf{C}_k\|\|\mathbf{b}^{(\ell-1)} - \mathbf{b}^*\|. \tag{69}$$

Therefore, similar to the convergence argument of (61) and (62), if $0 < \alpha < \frac{1}{\lambda_{\max}}$, we have $\|\mathbf{b}^{(\ell)} - \mathbf{b}^*\| \to 0$. This means that $\mathbf{B}^{(\ell)}$ converges to the optimal solution $\mathbf{B}^*$. The theorem is thus proved.

REFERENCES

[1] A. Cichocki and S. Amari, *Adaptive Blind Signal and Image Processing: Learning Algorithms and Applications*. New York, NY, USA: Wiley, 2002.

[2] R. Sameni, C. Jutten, and M. B. Shamsollahi, "A deflation procedure for subspace decomposition," *IEEE Trans. Signal Process.*, vol. 58, no. 4, pp. 2363–2374, Apr. 2010.

[3] R. Martín-Clemente, J. L. Camargo-Olivares, S. Hornillo-Mellado, M. Elena, and I. Román, "Fast technique for noninvasive fetal ECG extraction," *IEEE Trans. Biomed. Eng.*, vol. 58, no. 2, pp. 227–230, Feb. 2011.

[4] T. Tsalaile, R. Sameni, S. Sanei, C. Jutten, and J. Chambers, "Sequential blind source extraction for quasi-periodic signals with time-varying period," *IEEE Trans. Biomed. Eng.*, vol. 56, no. 3, pp. 646–655, Mar. 2009.

[5] P. Ahmadian, S. Sanei, L. Ascari, L. González-Villanueva, and M. A. Umiltà, "Constrained blind source extraction of readiness potentials from EEG," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 21, no. 4, pp. 567–575, Jul. 2013.

[6] Z. Koldovský, J. Málek, P. Tichavský, and F. Nesta, "Semi-blind noise extraction using partially known position of the target source," *IEEE/ACM Trans. Audio, Speech, Lang., Process.*, vol. 21, no. 10, pp. 2029–2041, Oct. 2013.

[7] R. Miyazaki *et al.*, "Musical-noise-free blind speech extraction integrating microphone array and iterative spectral subtraction," *Signal Process.*, vol. 102, pp. 226–239, Sep. 2014.

[8] F. D. Aprilyanti, J. Even, H. Saruwatari, K. Shikano, S. Nakamura, and T. Takatani, "Suppression of noise and late reverberation based on blind signal extraction and Wiener filtering," *Acoust. Sci. Technol.*, vol. 36, no. 4, pp. 302–313, Apr. 2015.

[9] D. Liu, S. Hu, and H. Zhang, "Simultaneous blind separation of instantaneous mixtures with arbitrary rank," *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 53, no. 10, pp. 2287–2298, Oct. 2006.

[10] A. Yeredor, "Non-orthogonal joint diagonalization in the least-squares sense with application in blind source separation," *IEEE Trans. Signal Process.*, vol. 50, no. 7, pp. 1545–1553, Jul. 2002.

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

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

[13] S. H. Sardouie, L. Albera, M. B. Shasollahi, and I. Merlet, "An efficient Jacobi-like deflationary ICA algorithm: Application to EEG denoising," *IEEE Signal Process. Lett.*, vol. 22, no. 8, pp. 1198–1202, Aug. 2015.

[14] M. Saleh, A. Karfoul, A. Kachenoura, I. Merlet, and L. Albera, "Efficient stepsize selection strategy for Givens parametrized ICA applied to EEG denoising," *IEEE Signal Process. Lett.*, vol. 24, no. 6, pp. 882–886, Aug. 2017.

[15] E. Ollila, "The deflation-based FastICA estimator: Statistical analysis revisited," *IEEE Trans. Signal Process.*, vol. 58, no. 3, pp. 1527–1541, Mar. 2010.

[16] S. C. Douglas, M. Gupta, H. Sawada, and S. Makino, "Spatio-temporal FastICA algorithms for the blind separation of convolutive mixtures," *IEEE Trans. Audio, Speech, Lang., Process.*, vol. 15, no. 5, pp. 1511–1520, Jul. 2007.

[17] L. T. Duarte, B. Rivet, and C. Jutten, "Blind extraction of smooth signals based on a second-order frequency identification algorithm," *IEEE Signal Process. Lett.*, vol. 17, no. 1, pp. 79–82, Jan. 2010.

[18] K. Zhang and L.-W. Chan, "Dimension reduction as a deflation method in ICA," *IEEE Signal Process. Lett.*, vol. 13, no. 1, pp. 45–48, Jan. 2006.

[19] V. Zarzoso and P. Comon, "Robust independent component analysis by iterative maximization of the kurtosis contrast with algebraic optimal step size," *IEEE Trans. Neural Netw.*, vol. 21, no. 2, pp. 248–261, Feb. 2010.

[20] A. Adib and D. Aboutajdine, "Reference-based blind source separation using a deflation approach," *Signal Process.*, vol. 85, no. 10, pp. 1943–1949, Oct. 2005.

[21] W. Liu, D. P. Mandic, and A. Cichocki, "Blind second-order source extraction of instantaneous noisy mixtures," *IEEE Trans. Circuits Syst. II, Exp. Briefs*, vol. 53, no. 9, pp. 931–935, Sep. 2006.

[22] Z. Shi and C. Zhang, "Blind source extraction using generalized autocorrelations," *IEEE Trans. Neural Netw.*, vol. 18, no. 5, pp. 1516–1524, Sep. 2007.

[23] B. Sällberg, N. Grbić, and I. Claesson, "Complex-valued independent component analysis for online blind speech extraction," *IEEE Trans. Audio, Speech, Lang., Process.*, vol. 16, no. 8, pp. 1624–1632, Nov. 2008.

[24] W. Y. Leong and D. P. Mandic, "Post-nonlinear blind extraction in the presence of ill-conditioned mixing," *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 55, no. 9, pp. 2631–2638, Oct. 2008.

[25] H. Sawada, S. Araki, R. Mukai, and S. Makino, "Blind extraction of dominant target sources using ICA and time-frequency masking," *IEEE Trans. Audio, Speech, Lang., Process.*, vol. 14, no. 6, pp. 2165–2173, Nov. 2006.

[26] H. Becker, L. Albera, P. Comon, A. Kachenoura, and I. Merlet, "A penalized semi-algebraic deflation ICA algorithm for the efficient extraction of interictal epileptic signals," *IEEE J. Biomed. Health Inform.*, vol. 21, no. 1, pp. 94–104, Jan. 2017.

[27] Y. Washizawa, Y. Yamashita, T. Tanaka, and A. Cichocki, "Blind extraction of global signal from multi-channel noisy observations," *IEEE Trans. Neural Netw.*, vol. 21, no. 9, pp. 1472–1481, Sep. 2010.

[28] J. Thomas, Y. Deville, and S. Hosseini, "Differential fast fixed-point algorithms for underdetermined instantaneous and convolutive partial blind source separation," *IEEE Trans. Signal Process.*, vol. 55, no. 7, pp. 3717–3729, Jul. 2007.

[29] A. Belouchrani, K. Abed-Meraim, J. F. Cardoso, and E. Moulines, "A blind source separation technique using second-order statistics," *IEEE Trans. Signal Process.*, vol. 45, no. 2, pp. 434–444, Feb. 1997.

[30] C.-F. Cardoso and A. Souloumiac, "Jacobi angles for simultaneous diagonalization," *SIAM J. Matrix Anal. Appl.*, vol. 17, no. 1, pp. 161–164, Jan. 1996.

[31] J. F. Cardoso and A. Souloumiac, "Blind beamforming for non-Gaussian signals," *Proc. Inst. Elect. Eng. F*, vol. 140, no. 6, pp. 362–370, Dec. 1993.

[32] E. Moreau, "A generalization of joint-diagonalization criteria for source separation," *IEEE Trans. Signal Process.*, vol. 49, no. 3, pp. 530–541, Mar. 2001.

[33] A. Ziehe, K. R. Muller, G. Nolte, B. M. Mackert, and G. Curio, "Artifact reduction in magnetoneurography based on time-delayed second-order correlations," *IEEE Trans. Biomed. Eng.*, vol. 47, no. 1, pp. 75–87, Jan. 2000.

[34] X. Li, X. Zhang, and P. Li, "Gradient search non-orthogonal approximate joint diagonalization algorithm," *Tsinghua Sci. Technol.*, vol. 12, no. 6, pp. 669–673, Dec. 2007.

[35] X. Luciani and L. Albera, "Joint eigenvalue decomposition of non-defective matrices based on the LU factorization with application to ICA," *IEEE Trans. Signal Process.*, vol. 63, no. 17, pp. 4594–4608, Sep. 2015.

[36] G. Zhou, S. Xie, Z. Yang, and J. Zhang, "Nonorthogonal approximate joint diagonalization with well-conditioned diagonalizers," *IEEE Trans. Neural Netw.*, vol. 20, no. 11, pp. 1810–1819, Nov. 2009.

[37] W. Fuxiang, L. Zhongkan, and Z. Jun, "A new joint diagonalization algorithm with application in blind source separation," *IEEE Signal Process. Lett.*, vol. 13, no. 1, pp. 41–44, Jan. 2006.

[38] X. L. Li and X. D. Zhang, "Nonorthogonal joint diagonalization free of degenerate solution," *IEEE Trans. Signal Process.*, vol. 55, no. 5, pp. 1803–1814, May 2007.

[39] G. Chabriel and J. Barrere, "A direct algorithm for nonorthogonal approximate joint diagonalization," *IEEE Trans. Signal Process.*, vol. 60, no. 1, pp. 39–47, Jan. 2012.

[40] A. Mesloub, K. Abed-Meraim, and A. Belouchrani, "A new algorithm for complex non-orthogonal joint diagonalization based on shear and givens rotations," *IEEE Trans. Signal Process.*, vol. 62, no. 8, pp. 1913–1925, Apr. 2014.

[41] S. Saito, K. Oishi, and T. Furukawa, "Convolutive blind source separation using an iterative least-squares algorithm for non-orthogonal approximate joint diagonalization," *IEEE/ACM Trans. Audio, Speech, Lang., Process.*, vol. 23, no. 12, pp. 1913–1925, Dec. 2015.

[42] A. Yeredor, "On using exact joint diagonalization for noniterative approximate joint diagonalization," *IEEE Signal Process. Lett.*, vol. 12, no. 9, pp. 645–648, Sep. 2005.

[43] L. Wang, L. Albera, A. Kachenoura, H. Shu, and L. Senhadji, "Nonnegative joint diagonalization by congruence based on LU matrix factorization," *IEEE Signal Process. Lett.*, vol. 20, no. 8, pp. 807–810, Aug. 2013.

[44] X. F. Xu, D. Z. Feng, and W. X. Zheng, "A fast algorithm for nonunitary joint diagonalization and its application to blind source separation," *IEEE Trans. Signal Process.*, vol. 59, no. 7, pp. 3457–3463, Jul. 2011.

[45] P. Tichavský and A. Yeredor, "Fast approximate joint diagonalization incorporating weight matrices," *IEEE Trans. Signal Process.*, vol. 57, no. 3, pp. 878–891, Mar. 2009.

[46] S. Degerine and E. Kane, "A comparative study of approximate joint diagonalization algorithms for blind source separation in presence of additive noise," *IEEE Trans. Signal Process.*, vol. 55, no. 6, pp. 3022–3031, Jun. 2007.

[47] G. Chabriel, J. Barrere, N. Thirion-Moreau, and E. Moreau, "Algebraic joint zero-diagonalization and blind sources separation," *IEEE Trans. Signal Process.*, vol. 56, no. 3, pp. 980–989, Mar. 2008.

[48] R. Vollgraf and K. Obermayer, "Quadratic optimization for simultaneous matrix diagonalization," *IEEE Trans. Signal Process.*, vol. 54, no. 9, pp. 3270–3278, Sep. 2006.

[49] L. D. Lathauwer, B. D. Moor, and J. Vandewalle, "Computation of the canonical decomposition by means of a simultaneous generalized Schur decomposition," *SIAM J. Matrix Anal. Appl.*, vol. 26, no. 2, pp. 295–327, Apr. 2004.

[50] Y.-B. Zhao, "An approximation theory of matrix rank minimization and its application to quadratic equations," *Linear Algebra Appl.*, vol. 59, no. 7, pp. 77–93, Jan. 2012.

[51] J. Tang, G. He, L. Dong, and L. Fang, "A smoothing Newton method for second-order cone optimization based on a new smoothing function," *Appl. Math. Comput.*, vol. 218, no. 4, pp. 1317–1329, 2011.

[52] J. Zhou, S.-C. Fang, and W. Xing, "Conic approximation to quadratic optimization with linear complementarity constraints," *Comput. Optim. Appl.*, vol. 66, pp. 97–122, 2017.

[53] M. Zangiabadi, G. Gu, and C. Roos, "A full Nesterov–Todd step infeasible interior-point method for second-order cone optimization," *J. Optim. Theory Appl.*, vol. 158, no. 3, pp. 816–858, 2013.

[54] Y. Hsia, S. Wang, and Z. Xu, "Improved semidefinite approximation bounds for nonconvex nonhomogeneous quadratic optimization with ellipsoid constraints," *Oper. Res. Lett.*, vol. 43, no. 4, pp. 378–383, 2015.

[55] A. Malcolm and D. P. Nicholls, "Operator expansions and constrained quadratic optimization for interface reconstruction: Impenetrable periodic acoustic media," *Wave Motion*, vol. 51, no. 1, pp. 23–40, Jan. 2014.

[56] M. González-Mendoza, R. E. I. Orozco, A. L. G. Gamboa, N. Hernández-Gress, J. Mora-Vargas, and J. C. López-Pimentel, "Quadratic optimization fine tuning for the support vector machines learning phase," *Expert Syst. Appl.*, vol. 41, no. 3, pp. 886–892, 2014.

[57] R. Oberdieck and E. N. Pistikopoulos, "Multi-objective optimization with convex quadratic cost functions: A multi-parametric programming approach," *Comput. Chem. Eng.*, vol. 85, pp. 36–39, 2016.

[58] S. Ma, D. Goldfarb, and L. Chen, "Fixed point and Bregman iterative methods for matrix rank minimization," *Math. Program. Ser. A*, vol. 128, no. 1, pp. 321–353, Jan. 2011.

[59] K. Mohan and M. Fazel, "Iterative reweighted algorithms for matrix rank minimization," *J. Mach. Learn. Res.*, vol. 13, no. 1, pp. 3441–3473, Nov. 2012.

[60] C. Lu, Z. Lin, and S. Yan, "Smoothed low rank and sparse matrix recovery by iteratively reweighted least squares minimization," *IEEE Trans. Image Process.*, vol. 24, no. 2, pp. 646–654, Feb. 2015.

[61] X. Lin and G. Wei, "Accelerated reweighted nuclear norm minimization algorithm for low rank matrix recovery," *Signal Process.*, vol. 114, pp. 24–33, 2015.

[62] J.-F. Cai, E. J. Candès, and Z. Shen, "A singular value thresholding algorithm for matrix completion," *SIAM J. Optim.*, vol. 20, no. 4, pp. 1956–1982, 2010.

[63] J. P. Haldar and D. Hernando, "Rank-constrained solutions to linear matrix equations using power factorization," *IEEE Signal Process. Lett.*, vol. 16, no. 7, pp. 584–587, Jul. 2009.

[64] K. Lee and Y. Bresler, "Admira: Atomic decomposition for minimum rank approximation," *IEEE Trans. Inf. Theory*, vol. 56, no. 9, pp. 4402–4416, Sep. 2010.

[65] P. Jain, R. Meka, and I. S. Dhillon, "Guaranteed rank minimization via singular value projection," in *Proc. Neural Inf. Process. Syst.*, 2010, pp. 937–945.

[66] K. Konishi, K. Uruma, T. Takahashi, and T. Furukawa, "Iterative partial matrix shrinkage algorithm for matrix rank minimization," *Signal Process.*, vol. 100, pp. 124–131, 2014.

[67] A. Hyvärinen, J. Karhunen, and E. Oja, *Independent Component Analysis*. New York, NY, USA: Wiley, 2001.

[68] P. Tichavský, A. E. Doron, and J. Nielsen, "A computationally affordable implementation of an asymptotically optimal BSS algorithm for AR sources," in *Proc. Eur. Signal Process. Conf.*, Florence, Italy, Sep. 4–8, 2006, pp. 1–5.

[69] P. Tichavský and A. Yeredor, "Fast approximate joint diagonalization incorporating weight matrices," *IEEE Trans. Signal Process.*, vol. 57, no. 3, pp. 878–891, Mar. 2009.

[70] P. Tichavský, "Matlab code for U-WEDGE, WEDGE, BG-WEDGE and WASOBI," 2009. [Online]. Available: http://si.utia.cas.cz/Tichavsky.html

[71] N. D. Sidiropoulos, "A short list of sound sources," 2010. [Online]. Available: http://www.telecom.tuc.gr/∼nikos/BSS_Nikos.html

[72] P. Tichavský and Z. Koldovský, "Fast and accurate methods of independent component analysis: A survey," *Kybernetika*, vol. 47, no. 3, pp. 426–438, 2011.
[73] G. H. Golub and C. F. V. Loan, *Matrix Computations*, 2nd ed. Baltimore, MD, USA: The Johns Hopkins Univ. Press, 1989.

Mingjian Zhang (M'17) received the B.E. and Ph.D. degrees from the South China University of Technology, Guangzhou, China, in 1998 and 2004, respectively.

He is currently an Associate Professor with the Department of Information Technology, Hunan Police Academy, China. He is also a Postdoctoral Researcher with the School of Information Science and Engineering, Central South University, Changsha, China. From 2005 to 2011, he was an Assistant, then an Associate Professor with the School of Automation, Guangdong University of Technology, China. From 2016 to 2017, he was with the Department of Electrical and Computer Engineering, State University of New York at Binghamton, USA, as a visiting scholar. His research interests include blind signal extraction, independent component analysis, machine learning, and speech signal processing.

Xiaohua Li (M'00–SM'06) received the B.S. and M.S. degrees from Shanghai Jiao Tong University, Shanghai, China, in 1992 and 1995, respectively, and the Ph.D. degree in electrical engineering from the University of Cincinnati, Cincinnati, OH, USA, in 2000. He was an Assistant Professor from 2000 to 2006, and has been an Associate Professor since 2006, both with the Department of Electrical and Computer Engineering, State University of New York at Binghamton, Binghamton, NY, USA. His research interests include the fields of signal processing, machine learning, wireless communications, and wireless information assurance. He was the recipient of the Best Reviewer Award in 2013 from the *ETRI Journal*, and the Best Paper Award of the 7th International Conference on Mathematical Methods, Models and Architectures for Computer Networks Security (MMM-ACNS'2017).

Jun Peng (M'08) received the B.S. degree from Xiangtan University, Xiangtan, China, in 1987, the M.Sc. degree from the National University of Defense Technology, Changsha, China, in 1990, and the Ph.D. degree from Central South University, Changsha, China, in 2005. She is currently a Professor with the School of Information Science and Engineering, Central South University. Her research interests include cooperative control, cloud computing, and wireless communications. In April 1990, she joined the Staff of Central South University. From 2006 to 2007, she was with the School of Electrical and Computer Science, University of Central Florida, USA, as a visiting scholar.