Received January 13, 2017, accepted January 31, 2017, date of publication February 16, 2017, date of current version March 13, 2017.

Digital Object Identifier 10.1109/ACCESS.2017.2664978

Second-Order Cyclostationary Statistics-Based Blind Source Extraction From Convolutional Mixtures

YONG XIANG[1], (Senior Member, IEEE), DEZHONG PENG[2], (Member, IEEE), INDIVARIE UBHAYARATNE[1], BERNARD ROLFE[3], AND MICHAEL PEREIRA[3]
[1]School of Information Technology, Deakin University, Burwood campus, VIC 3125, Australia
[2]Machine Intelligence Laboratory, College of Computer Science, Sichuan University, Chengdu 610065, China
[3]School of Engineering, Deakin University, Waurn Ponds campus, VIC 3220, Australia

Corresponding author: Y. Xiang (yxiang@deakin.edu.au)

This work was supported in part by the Australian Research Council under Grant DP110102L076 and Grant LP120100239 and in part by the National Natural Science Foundation of China under Grant 61432012 and Grant 61172180.

ABSTRACT Blind source extraction (BSE) aims to extract the source of interest (SOI) from the outputs of a mixing system, which is a challenging problem. A property existing in many signals is cyclostationarity and this property has been widely exploited in BSE. While various cyclostationarity-based BSE methods have been reported in the literature, they usually require the mixing system to be instantaneous. In this paper, we address BSE in the context that the mixing system is convolutional. Specifically, a new BSE method is developed to extract cyclostationary source signal from the outputs of a multiple-input-multiple-output finite-impulse-response mixing system. It is shown that if the SOI has a unique cyclostationary frequency, it can be recovered from the measured data. The effectiveness of the proposed BSE method is demonstrated by simulation results.

INDEX TERMS Blind signal extraction, cyclostationary signal, MIMO FIR mixing system, second-order cyclostationary statistics.

I. INTRODUCTION

Blind source separation (BSS) is a fundamental problem encountered in many applications [1]–[4]. It considers the practical multiple-input-multiple-output (MIMO) scenarios and aims to estimate the unknown source signals from the measured data by sensors. To achieve BSS, the sources should have some sort of differences, e.g., they are mutually independent/uncorrelated [1], cyclostationary [5], of constant modulus [3], [6], etc. In some man-made systems such as wireless communication systems, the source signals can be pre-processed such that they possess certain diversity [4]. Among different diversity properties, cyclostationarity exists in many practical applications such as communication, telemetry, radar, sonar and mechanical systems. For example, cyclostationarity is shared by most communication signals as a result of periodic switching, gating, or mixing operations at the transmitter [7]. Many signals generated by mechanical systems are also cyclostationary due to the periodical rotation or movement of some parts [2].

Among the existing cyclostationarity based BSS methods, most of them require the mixing system to be instantaneous. Some representative cyclostationarity based BSS methods for instantaneous mixing systems can be found in [5] and [8]–[14]. Although instantaneous mixing systems exist in practice, the mixing systems encountered in most real-world applications are actually dynamic, due to multi-path propagation [15]–[18]. Hence, in this paper, we relax the assumption on the mixing systems to allow them to be convolutional. Specifically, we assume that the MIMO mixing systems are of finite impulse response (FIR). So far, some cyclostationarity based BSS methods for MIMO FIR mixing systems have been developed but they are far from mature. The methods in [15]–[17] use nonlinear cost functions exploiting the higher-order cyclostationary statistics (HOCS) of the measured data, from which several adaptive and iterative algorithms have been derived. However, these algorithms are not globally convergent. Moreover, it is known that the HOCS-based methods normally require a larger number of data samples to achieve good statistical performance. This would significantly increase computational cost.

In contrast, the methods based on second-order cyclostationary statistics (SOCS) are often more efficient in computation. In [18], a frequency domain method is proposed to separate cyclostationary signals from MIMO FIR mixtures, which only employs the SOCS of the measured data. However, it needs to perform approximate joint diagonalization of a set of spectral correlation density matrices in each frequency bin, which introduces a frequency-dependent permutation ambiguity. To correct such permutation ambiguity, the source signals must satisfy some conditions: they are real-valued and mutually uncorrelated, and have different cyclic frequencies. If one or more of these conditions do not hold, significant performance degradation will be unavoidable. Moreover, this correction procedure itself will add additional errors to the final source separation outcome.

It should be noted that in some practical applications, one only needs to extract the source of interest (SOI), not all source signals, from the measured data [19]–[21]. This is a special case of BSS, which is named as blind source extraction (BSE). In [19]–[21], some BSE methods have been developed to extract a cyclostationary signal, which is of interest, from the outputs of an instantaneous mixing system. The method in [22] can deal with MIMO FIR mixing systems but it relies on the HOCS of the system outputs. In [23], an SOCS-based frequency domain method is proposed for BSE from MIMO FIR mixtures. Since this method performs BSE in the frequency domain, it utilizes cyclic power spectrum instead of cyclic correlation function. As a result, the additive noise term cannot be removed when estimating the extraction vector, which reduces the accuracy of signal extraction. Another SOCS-based BSE method for MIMO FIR mixing systems is reported in [24]. However, it depends on the assumption that the entries of the first column of the mixing matrix are constants instead of polynomials. This implies that the sensors should be placed close to the SOI, which would be restrictive in practice. Moreover, it assumes that for the SOI, the ratio of its cyclic correlation function over its correlation function is known. This assumption is impractical. Furthermore, its validity is only illustrated by preliminary simulation results.

In this paper, we propose a new SOCS-based BSE method for MIMO FIR mixing systems. The proposed method allows all of the entries of the mixing matrix to be polynomials, which relaxes the restriction on sensor placement. In addition, the new method does not require any prior knowledge of the ratio of the SOI's cyclic correlation function over its correlation function, as it has an inherent mechanism to calculate this ratio. The effectiveness of our method is verified both theoretically and experimentally in detail.

The remainder of the paper is as follows. Section II introduces the problem of BSE together with relevant assumptions. The new SOCS-based BSE method is proposed in Section III. Simulation results are provided in Section IV to illustrate the performance of the proposed method. Section V concludes the paper.

II. PROBLEM FORMULATION

The mixing system we consider is as follows:

$$\mathbf{y}(n) = \mathbf{H}(z)[\mathbf{x}(n)] + \mathbf{w}(n) \tag{1}$$

where $\mathbf{x}(n) = [x_1(n), x_2(n), \ldots, x_I(n)]^T$ is the source signal vector, $\mathbf{y}(n) = [y_1(n), y_2(n), \ldots, y_J(n)]^T$ is the system output vector, $\mathbf{w}(n) = [w_1(n), w_2(n), \ldots, w_J(n)]^T$ is the additive noise vector, $\mathbf{H}(z)$ is the $J \times I$ FIR mixing matrix, $\mathbf{H}(z)[\mathbf{x}(n)]$ denotes the operation of $\mathbf{H}(z)$ on $\mathbf{x}(n)$, and the superscript ^{T} stands for transpose. Here, $\mathbf{x}(n)$, $\mathbf{y}(n)$, $\mathbf{w}(n)$ and the coefficients of the polynomials in $\mathbf{H}(z)$ can be either real or complex valued. Without loss of generality, we consider $x_1(n)$ as the SOI. As in [24], we also assume that

A1) The SOI, $x_1(n)$, is zero-mean, temporally-white, cyclostationary with a unique nonzero cyclic frequency β_1, and independent of the other $I-1$ source signals $x_2(n)$, $x_3(n), \ldots, x_I(n)$.

A2) The mixing matrix $\mathbf{H}(z)$ has more rows than columns, i.e., $J > I$, and is irreducible and column-reduced.

A3) The noise signals $w_1(n), w_2(n), \ldots, w_J(n)$ are mutually independent with zero mean and equal variance σ_w^2. They are also independent of the source signals.

From the assumptions A1) and A3), it follows:

$$\langle x_1(n)x_1^*(n)e^{\jmath\beta n}\rangle > 0, \quad \text{if } \beta = 0, \beta_1 \tag{2}$$

$$\langle x_1(n)x_1^*(n-\tau)e^{\jmath\beta n}\rangle = 0, \quad \text{if } \beta = 0, \beta_1 but \quad \tau \neq 0 \tag{3}$$

$$\langle x_i(n)x_i^*(n)e^{\jmath\beta_1 n}\rangle = 0, \quad if\ i \neq 1 \tag{4}$$

$$\langle x_1(n)x_i^*(n)e^{\jmath\beta n}\rangle = 0, \quad if\ \beta = 0, \beta_1 \text{ but } i \neq 1 \tag{5}$$

$$\langle x_i(n)x_1^*(n)e^{\jmath\beta n}\rangle = 0, \quad if\ \beta = 0, \beta_1 \text{ but } i \neq 1 \tag{6}$$

$$\left\langle x_i(n)w_j^*(n)e^{\jmath\beta_1 n}\right\rangle = 0, \quad \forall i, j \tag{7}$$

$$\left\langle w_i(n)w_j^*(n)e^{\jmath\beta_1 n}\right\rangle = 0, \quad \forall i, j \tag{8}$$

where $\jmath = \sqrt{-1}$, the superscript * is the complex conjugate operator, and $< \cdot >$ denotes the time averaging operator defined as

$$\langle x_i(n)x_i^*(n-\tau)e^{\jmath\beta_1 n}\rangle = \lim_{N\to\infty} \frac{1}{N} \sum_{n=0}^{N-1} x_i(n)x_i^*(n-\tau)e^{\jmath\beta_1 n}$$

where N is the number of samples. The objective of BSE is to recover the cyclostationary SOI, $x_1(n)$, from the mixtures $y_1(n), y_2(n), \ldots, y_J(n)$, or equivalently $\mathbf{y}(n)$.

III. PROPOSED METHOD

In order to develop an algorithm to extract the cyclostationary signal $x_1(n)$ from the mixtures $\mathbf{y}(n)$, we first introduce the sliding model of the mixing system described by (1).

A. SLIDING MODEL OF THE MIXING SYSTEM

Let

$$\mathbf{H}(z) = \sum_{l=0}^{L} \mathbf{H}(l) z^{-l} \tag{9}$$

where L is the order of the MIMO FIR mixing matrix $\mathbf{H}(z)$. Denoting the order of the jth column of $\mathbf{H}(z)$ by L_j, it follows

$$L = max(L_1, L_2, \ldots, L_I). \tag{10}$$

Based on the ith system output $y_i(n)$, we define

$$\tilde{\mathbf{y}}_i(n) = [y_i(n), y_i(n-1), \ldots, y_i(n-W+1)]^T \tag{11}$$

where the slide-window width W is chosen to satisfy

$$W > \bar{L} = \sum_{i=1}^{I} L_i. \tag{12}$$

The purpose of ensuring the inequality in (12) will become clear in the next paragraph. It is worth noting that selecting a suitable W does not require information of the exact values of $L_1, L_2, \ldots, L_I$ but their upper bound L_{upper},[1] where

$$L_{upper} \geq L_1, L_2, \ldots, L_I. \tag{13}$$

We can simply choose

$$W > IL_{upper} \tag{14}$$

which apparently satisfies (12).

Denote the (i,j)th entry of $\mathbf{H}(l)$ by $h_{i,j}(l)$. Then, it follows from (1), (9) and (11) that

$$\tilde{\mathbf{y}}_i(n) = \sum_{j=1}^{I} \boldsymbol{\mathcal{H}}_{i,j}\tilde{\mathbf{x}}_j(n) + \tilde{\mathbf{w}}_i(n)$$

where

$$\tilde{\mathbf{x}}_j(n) = [x_j(n), x_j(n-1), \ldots, x_j(n-W-L_j+1)]^T \tag{15}$$

$$\tilde{\mathbf{w}}_i(n) = [w_i(n), w_i(n-1), \ldots, w_i(n-W+1)]^T \tag{16}$$

and $\boldsymbol{\mathcal{H}}_{i,j}$ is shown at the bottom of this page. Further denoting

$$\tilde{\mathbf{y}}(n) = [\tilde{\mathbf{y}}_1(n), \tilde{\mathbf{y}}_2(n), \ldots, \tilde{\mathbf{y}}_J(n)] \tag{17}$$

we have

$$\tilde{\mathbf{y}}(n) = \boldsymbol{\mathcal{H}}\tilde{\mathbf{x}}(n) + \tilde{\mathbf{w}}(n) \tag{18}$$

where

$$\tilde{\mathbf{x}}(n) = [\tilde{\mathbf{x}}_1(n), \tilde{\mathbf{x}}_2(n), \ldots, \tilde{\mathbf{x}}_I(n)] \tag{19}$$

$$\tilde{\mathbf{w}}(n) = [\tilde{\mathbf{w}}_1(n), \tilde{\mathbf{w}}_2(n), \ldots, \tilde{\mathbf{w}}_J(n)] \tag{20}$$

and

$$\boldsymbol{\mathcal{H}} = \begin{bmatrix} \boldsymbol{\mathcal{H}}_{1,1} & \boldsymbol{\mathcal{H}}_{1,2} & \cdots & \boldsymbol{\mathcal{H}}_{1,I} \\ \boldsymbol{\mathcal{H}}_{2,1} & \boldsymbol{\mathcal{H}}_{2,2} & \cdots & \boldsymbol{\mathcal{H}}_{2,I} \\ \vdots & \vdots & \vdots & \vdots \\ \boldsymbol{\mathcal{H}}_{J,1} & \boldsymbol{\mathcal{H}}_{J,2} & \cdots & \boldsymbol{\mathcal{H}}_{J,I} \end{bmatrix}_{JW\times(IW+\bar{L})}. \tag{21}$$

Since $J > I$, the matrix $\boldsymbol{\mathcal{H}}$ is a tall matrix, i.e., it has more rows than columns. Also, under the assumption A1) and the inequality in (12), it is shown in [25] that the matrix $\boldsymbol{\mathcal{H}}$ has the full column rank $IW + \bar{L}$.

[1] Depending on the application scenario, L_{upper} can be chosen empirically or by experiment.

Let $\mathbf{u}$ be a $JW \times 1$ vector and define the compound vector

$$\mathbf{c} = \left(\mathbf{u}^H\boldsymbol{\mathcal{H}}\right)^H \tag{22}$$

where the superscript ^{H} stands for complex conjugate transpose. Then,

$$\begin{aligned} \mathbf{u}^H\tilde{\mathbf{y}}(n) &= \mathbf{u}^H\boldsymbol{\mathcal{H}}\tilde{\mathbf{x}}(n) + \mathbf{u}^H\tilde{\mathbf{w}}(n) \\ &= \mathbf{c}^H\tilde{\mathbf{x}}(n) + \mathbf{u}^H\tilde{\mathbf{w}}(n). \end{aligned}$$

Clearly, the first $W + L_1$ elements of $\mathbf{c}$ correspond to the first source signal $x_1(n)$ and its delayed versions $x_1(n-1)$, $x_1(n-2), \cdots, x_1(n-W-L_1+1)$, respectively. So, to recover $x_1(n)$ from $\tilde{\mathbf{y}}(n)$, we need to find such a vector $\mathbf{u}$ that the first element of $\mathbf{c}$ is nonzero and the rest elements of $\mathbf{c}$ are zero.

B. ALGORITHM DEVELOPMENT

In order to recover $x_1(n)$ from $\tilde{\mathbf{y}}(n)$, it is essential to taking advantage of the properties of the sources. So, to begin with, we look into the features of some source correlation matrices.

1) FEATURES OF SOURCE CORRELATION MATRICES

Define the cyclic correlation function $\rho_{ij}^{\beta_1}(\tau)$ as

$$\rho_{ij}^{\beta_1}(\tau) = \left\langle x_i(n)x_j^*(n-\tau)e^{\jmath\beta_1 n}\right\rangle \tag{23}$$

and the cyclic correlation matrix $\mathbf{R}_{\tilde{\mathbf{x}}_i\tilde{\mathbf{x}}_j}^{\beta_1}(\tau)$ as

$$\mathbf{R}_{\tilde{\mathbf{x}}_i\tilde{\mathbf{x}}_j}^{\beta_1}(\tau) = \left\langle \tilde{\mathbf{x}}_i(n)\tilde{\mathbf{x}}_j^H(n-\tau)e^{\jmath\beta_1 n}\right\rangle. \tag{24}$$

Let $\mathbf{I}_i$ be the $i \times i$ identity matrix and $\mathbf{J}_i$ stand for the $i \times i$ Jordan matrix with the following form:

$$\mathbf{J}_i = \begin{bmatrix} 0 & 0 & \cdots & \cdots & 0 \\ 1 & 0 & \ddots & \ddots & \vdots \\ 0 & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & 1 & 0 \end{bmatrix}. \tag{25}$$

From (2)-(8), (15), (16) and (23)-(25), we obtain

$$\begin{aligned} \mathbf{R}_{\tilde{\mathbf{x}}_1\tilde{\mathbf{x}}_1}^{\beta_1}(1) &= \begin{bmatrix} 0 & 0 & \cdots & \cdots & 0 \\ \rho_{11}^{\beta_1}(0) & 0 & \ddots & \ddots & \vdots \\ 0 & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \rho_{11}^{\beta_1}(0) & 0 \end{bmatrix} \\ &= \rho_{11}^{\beta_1}(0) \cdot \mathbf{J}_{W+L_1} \end{aligned} \tag{26}$$

$$\boldsymbol{\mathcal{H}}_{i,j} = \begin{bmatrix} h_{i,j}(0) & \cdots & \cdots & h_{i,j}(L_j) & 0 & \cdots & 0 \\ 0 & h_{i,j}(0) & \cdots & \cdots & h_{i,j}(L_j) & \cdots & 0 \\ & & \ddots & & & \ddots & \\ 0 & \cdots & 0 & h_{i,j}(0) & \cdots & \cdots & h_{i,j}(L_j) \end{bmatrix}_{W\times(W+L_j)}$$

$$\mathbf{R}^{\beta_1}_{\tilde{\mathbf{x}}_i\tilde{\mathbf{x}}_i}(\tau) = \mathbf{0}, \quad \text{if } i \neq 1 \tag{27}$$
$$\mathbf{R}^{\beta_1}_{\tilde{\mathbf{x}}_i\tilde{\mathbf{x}}_j}(\tau) = \mathbf{0}, \quad \text{if } i \neq j \tag{28}$$
$$\mathbf{R}^{\beta_1}_{\tilde{\mathbf{x}}_i\tilde{\mathbf{w}}_j}(\tau) = \mathbf{0}, \quad \forall i, j \tag{29}$$
$$\mathbf{R}^{\beta_1}_{\tilde{\mathbf{w}}_i\tilde{\mathbf{w}}_j}(\tau) = \mathbf{0}, \quad \forall i, j. \tag{30}$$

Besides, we define the correlation function $\rho_{ij}(\tau)$ as

$$\rho_{ij}(\tau) = \left\langle x_i(n)x_j^*(n-\tau)\right\rangle \tag{31}$$

and the correlation matrix $\mathbf{R}_{\tilde{\mathbf{x}}_i\tilde{\mathbf{x}}_j}(\tau)$ as

$$\mathbf{R}_{\tilde{\mathbf{x}}_i\tilde{\mathbf{x}}_j}(\tau) = \left\langle \tilde{\mathbf{x}}_i(n)\tilde{\mathbf{x}}_j^H(n-\tau)\right\rangle. \tag{32}$$

It follows that

$$\mathbf{R}_{\tilde{\mathbf{x}}_1\tilde{\mathbf{x}}_1}(0) = \begin{bmatrix} \rho_{ii}(0) & 0 & 0 & \cdots & 0 & 0 \\ 0 & \rho_{ii}(0) & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 0 & \rho_{ii}(0) \end{bmatrix} = \rho_{ii}(0)\cdot \mathbf{I}_{W+L_1} \tag{33}$$

and

$$\mathbf{R}_{\tilde{\mathbf{x}}_1\tilde{\mathbf{x}}_i}(\tau) = \mathbf{R}_{\tilde{\mathbf{x}}_i\tilde{\mathbf{x}}_1}(\tau) = \mathbf{0}, \quad \text{if } i \neq 1. \tag{34}$$

Next, we will construct a set of output correlation matrices from the available system outputs. Due to the features of the source correlation matrices, these output correlation matrices have some special structures.

2) CONSTRUCTION OF OUTPUT CORRELATION MATRICES

Similar to (24), we construct the cyclic correlation matrix of $\tilde{\mathbf{y}}(n)$ at time lag τ by

$$\mathbf{R}^{\beta_1}_{\tilde{\mathbf{y}}\tilde{\mathbf{y}}}(\tau) = \left\langle \tilde{\mathbf{y}}(n)\tilde{\mathbf{y}}^H(n-\tau)e^{\jmath\beta_1 n}\right\rangle. \tag{35}$$

We have the following lemma.

Lemma 1: The rank of $\mathbf{R}^{\beta_1}_{\tilde{\mathbf{y}}\tilde{\mathbf{y}}}(1)$ is $W + L_1 - 1$.

Proof: From (17)-(21), (26)-(30) and (35), we have

$$\begin{aligned}\mathbf{R}^{\beta_1}_{\tilde{\mathbf{y}}\tilde{\mathbf{y}}}(1) &= \left\langle \tilde{\mathbf{y}}(n)\tilde{\mathbf{y}}^H(n-1)e^{\jmath\beta_1 n}\right\rangle \\ &= \mathcal{H}\mathbf{R}^{\beta_1}_{\tilde{\mathbf{x}}\tilde{\mathbf{x}}}(1)\mathcal{H}^H + \mathcal{H}\mathbf{R}^{\beta_1}_{\tilde{\mathbf{x}}\tilde{\mathbf{w}}}(1) \\ &\quad + (\mathcal{H}\mathbf{R}^{\beta_1}_{\tilde{\mathbf{x}}\tilde{\mathbf{w}}}(1))^H + \mathbf{R}^{\beta_1}_{\tilde{\mathbf{w}}\tilde{\mathbf{w}}}(1) \\ &= \mathcal{H}\mathbf{R}^{\beta_1}_{\tilde{\mathbf{x}}\tilde{\mathbf{x}}}(1)\mathcal{H}^H \\ &= \mathcal{H}\begin{bmatrix} \rho^{\beta_1}_{11}(0)\mathbf{J}_{W+L_1} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} \\ \vdots & \vdots & \vdots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} \end{bmatrix}\mathcal{H}^H. \end{aligned} \tag{36}$$

Since the $(W+L_1)\times(W+L_1)$ matrix $\mathbf{J}_{W+L_1}$ is rank deficient by 1 and $\mathcal{H}$ is of full column rank, the rank of $\mathbf{R}^{\beta_1}_{\tilde{\mathbf{y}}\tilde{\mathbf{y}}}(1)$ is $W + L_1 - 1$. This completes the proof. □

Based on Lemma 1, $\mathbf{R}^{\beta_1}_{\tilde{\mathbf{y}}\tilde{\mathbf{y}}}(1)$ has $W + L_1 - 1$ nonzero singular values. Suppose that $\mathbf{u}_{1,1}, \mathbf{u}_{1,2}, \cdots, \mathbf{u}_{1,K_1}$ are the singular vectors corresponding to the K_1 largest singular values of $\mathbf{R}^{\beta_1}_{\tilde{\mathbf{y}}\tilde{\mathbf{y}}}(1)$, where

$$K_1 = W + L_{upper} - 1 \geq W + L_1 - 1. \tag{37}$$

We can form the following matrix

$$\mathbf{U}_1 = \left[\mathbf{u}_{1,1}, \mathbf{u}_{1,2}, \ldots, \mathbf{u}_{1,K_1}\right]. \tag{38}$$

As will be shown later, $\mathbf{U}_1$ plays an important role in the derivation of the desired SOI extraction vector.

Similar to (36), we can also find

$$\mathbf{R}^{\beta_1}_{\tilde{\mathbf{y}}\tilde{\mathbf{y}}}(0) = \mathcal{H}\begin{bmatrix} \rho^{\beta_1}_{11}(0)\mathbf{I}_{W+L_1} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} \\ \vdots & \vdots & \vdots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} \end{bmatrix}\mathcal{H}^H. \tag{39}$$

Furthermore, from (17)-(21) and (31)-(34), the autocorrelation matrix of $\tilde{\mathbf{y}}(n)$ can be obtained as follows:

$$\begin{aligned}\mathbf{R}_{\tilde{\mathbf{y}}\tilde{\mathbf{y}}}(0) &= \left\langle \tilde{\mathbf{y}}(n)\tilde{\mathbf{y}}^H(n)\right\rangle \\ &= \mathcal{H}\mathbf{R}_{\tilde{\mathbf{x}}\tilde{\mathbf{x}}}(0)\mathcal{H}^H + \mathcal{H}\mathbf{R}_{\tilde{\mathbf{x}}\tilde{\mathbf{w}}}(0) \\ &\quad + (\mathcal{H}\mathbf{R}_{\tilde{\mathbf{x}}\tilde{\mathbf{w}}}(0))^H + \mathbf{R}_{\tilde{\mathbf{w}}\tilde{\mathbf{w}}}(0) \\ &= \mathcal{H}\mathbf{R}_{\tilde{\mathbf{x}}\tilde{\mathbf{x}}}(0)\mathcal{H}^H + \mathbf{R}_{\tilde{\mathbf{w}}\tilde{\mathbf{w}}}(0) \\ &= \mathcal{H}\mathbf{R}_{\tilde{\mathbf{x}}\tilde{\mathbf{x}}}(0)\mathcal{H}^H + \sigma_w^2\mathbf{I}_{JW}. \end{aligned} \tag{40}$$

Since $\mathcal{H}$ is a tall matrix, (40) implies that σ_w^2 is the smallest eigenvalue of $\mathbf{R}_{\tilde{\mathbf{y}}\tilde{\mathbf{y}}}(0)$ and thus can be estimated from $\mathbf{R}_{\tilde{\mathbf{y}}\tilde{\mathbf{y}}}(0)$. Consequently, we can remove $\sigma_w^2\mathbf{I}_{JW}$ from $\mathbf{R}_{\tilde{\mathbf{y}}\tilde{\mathbf{y}}}(0)$ by

$$\bar{\mathbf{R}}_{\tilde{\mathbf{y}}\tilde{\mathbf{y}}}(0) = \mathbf{R}_{\tilde{\mathbf{y}}\tilde{\mathbf{y}}}(0) - \sigma_w^2\mathbf{I}_{JW} \tag{41}$$

which leads to

$$\begin{aligned}\bar{\mathbf{R}}_{\tilde{\mathbf{y}}\tilde{\mathbf{y}}}(0) &= \mathcal{H}\mathbf{R}_{\tilde{\mathbf{x}}\tilde{\mathbf{x}}}(0)\mathcal{H}^H \\ &= \mathcal{H}\begin{bmatrix} \mathbf{R}_{\tilde{\mathbf{x}}_1\tilde{\mathbf{x}}_1}(0) & \mathbf{R}_{\tilde{\mathbf{x}}_1\tilde{\mathbf{x}}_2}(0) & \cdots & \mathbf{R}_{\tilde{\mathbf{x}}_1\tilde{\mathbf{x}}_J}(0) \\ \mathbf{R}_{\tilde{\mathbf{x}}_2\tilde{\mathbf{x}}_1}(0) & \mathbf{R}_{\tilde{\mathbf{x}}_2\tilde{\mathbf{x}}_2}(0) & \cdots & \mathbf{R}_{\tilde{\mathbf{x}}_2\tilde{\mathbf{x}}_J}(0) \\ \vdots & \vdots & \vdots & \vdots \\ \mathbf{R}_{\tilde{\mathbf{x}}_J\tilde{\mathbf{x}}_1}(0) & \mathbf{R}_{\tilde{\mathbf{x}}_J\tilde{\mathbf{x}}_2}(0) & \cdots & \mathbf{R}_{\tilde{\mathbf{x}}_J\tilde{\mathbf{x}}_J}(0) \end{bmatrix}\mathcal{H}^H. \end{aligned} \tag{42}$$

Based on $\mathbf{R}^{\beta_1}_{\tilde{\mathbf{y}}\tilde{\mathbf{y}}}(0)$ and $\bar{\mathbf{R}}_{\tilde{\mathbf{y}}\tilde{\mathbf{y}}}(0)$, we introduce a $JW \times JW$ matrix $\mathbf{T}(\xi)$ defined by

$$\mathbf{T}(\xi) = \mathbf{R}^{\beta_1}_{\tilde{\mathbf{y}}\tilde{\mathbf{y}}}(0) - \xi\cdot\bar{\mathbf{R}}_{\tilde{\mathbf{y}}\tilde{\mathbf{y}}}(0) \tag{43}$$

where ξ is a nonzero constant. In the next subsection, we will show how to construct the matrix $\mathbf{T}(\xi)$ in a meaningful way.

3) CONSTRUCTION OF $\mathbf{T}(\xi)$

From (43), it is clear that if ξ is given, $\mathbf{T}(\xi)$ can be easily constructed by using $\mathbf{R}^{\beta_1}_{\tilde{\mathbf{y}}\tilde{\mathbf{y}}}(0)$ and $\bar{\mathbf{R}}_{\tilde{\mathbf{y}}\tilde{\mathbf{y}}}(0)$. To find the right value for ξ, it is interesting to see how the value of ξ affects $\mathbf{T}(\xi)$, which is shown in the lemma below.

Lemma 2: The matrix $\mathbf{T}(\xi)$ is rank deficient. Its rank is $IW + \bar{L}$ if and only if

$$\xi \neq \rho^{\beta_1}_{11}(0)/\rho_{11}(0) \tag{44}$$

and it reduces to $(I-1)W + \bar{L} - L_1$ if and only if

$$\xi = \rho^{\beta_1}_{11}(0)/\rho_{11}(0). \tag{45}$$

Proof: From (33), (34) and (42), it follows

$$\begin{aligned}
&\bar{\mathbf{R}}_{\tilde{\mathbf{y}}\tilde{\mathbf{y}}}(0)\\
&= \boldsymbol{\mathcal{H}} \begin{bmatrix} \rho_{11}(0)\cdot \mathbf{I}_{W+L_1} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{R}_{\tilde{\mathbf{x}}_2\tilde{\mathbf{x}}_2}(0) & \cdots & \mathbf{R}_{\tilde{\mathbf{x}}_2\tilde{\mathbf{x}}_J}(0) \\ \vdots & \vdots & \vdots & \vdots \\ \mathbf{0} & \mathbf{R}_{\tilde{\mathbf{x}}_J\tilde{\mathbf{x}}_2}(0) & \cdots & \mathbf{R}_{\tilde{\mathbf{x}}_J\tilde{\mathbf{x}}_J}(0) \end{bmatrix} \boldsymbol{\mathcal{H}}^H\\
&= \boldsymbol{\mathcal{H}} \begin{bmatrix} \rho_{11}(0)\cdot \mathbf{I}_{W+L_1} & \mathbf{0} \\ \mathbf{0} & \bar{\mathbf{R}}_{\tilde{\mathbf{x}}\tilde{\mathbf{x}}}(0) \end{bmatrix} \boldsymbol{\mathcal{H}}^H \qquad (46)
\end{aligned}$$

where

$$\bar{\mathbf{R}}_{\tilde{\mathbf{x}}\tilde{\mathbf{x}}}(0) = \begin{bmatrix} \mathbf{R}_{\tilde{\mathbf{x}}_2\tilde{\mathbf{x}}_2}(0) & \cdots & \mathbf{R}_{\tilde{\mathbf{x}}_2\tilde{\mathbf{x}}_J}(0) \\ \vdots & \vdots & \vdots \\ \mathbf{R}_{\tilde{\mathbf{x}}_J\tilde{\mathbf{x}}_2}(0) & \cdots & \mathbf{R}_{\tilde{\mathbf{x}}_J\tilde{\mathbf{x}}_J}(0) \end{bmatrix}$$

is of full rank. Based on (39) and (46), we can rewrite (43) as

$$\begin{aligned}
&\mathbf{T}(\xi)\\
&= \boldsymbol{\mathcal{H}} \begin{bmatrix} \left(\rho_{11}^{\beta_1}(0) - \xi\cdot\rho_{11}(0)\right)\mathbf{I}_{W+L_1} & \mathbf{0} \\ \mathbf{0} & -\xi\cdot\bar{\mathbf{R}}_{\tilde{\mathbf{x}}\tilde{\mathbf{x}}}(0) \end{bmatrix} \boldsymbol{\mathcal{H}}^H. \qquad (47)
\end{aligned}$$

Since $\boldsymbol{\mathcal{H}}$ is a full column rank tall matrix, one can see from (47) that $\mathbf{T}(\xi)$ is rank deficient. Moreover, since the dimension of $\boldsymbol{\mathcal{H}}$ is $JW \times (IW+\bar{L})$, it is obvious from (47) that the rank of $\mathbf{T}(\xi)$ is $IW + \bar{L}$ if and only if $\xi \neq \rho_{11}^{\beta_1}(0)/\rho_{11}(0)$, and it reduces to $(I-1)W + \bar{L} - L_1$ if and only if $\xi = \rho_{11}^{\beta_1}(0)/\rho_{11}(0)$. This completes the proof. □

As will become clear in the next subsection, the value of ξ satisfying (45) is desirable. However, ξ cannot be directly computed using (45) as $\rho_{11}^{\beta_1}(0)$ and $\rho_{11}(0)$ are often unknown in practical applications. Now, we show that ξ can be estimated by utilizing $\mathbf{R}_{\tilde{\mathbf{y}}\tilde{\mathbf{y}}}^{\beta_1}(0)$ and $\bar{\mathbf{R}}_{\tilde{\mathbf{y}}\tilde{\mathbf{y}}}(0)$.

First, before ξ is determined, $\mathbf{T}(\xi)$ in (43) can be viewed as a symbolic matrix with the variable ξ. Let us denote the rth singular value of $\mathbf{T}(\xi)$ by $E_r(\xi)$. From Lemma 2, it implies that

$$E_r(\xi) = 0, \quad \text{for } (I-1)W + \bar{L} - L_1 < r < IW + \bar{L} + 1 \qquad (48)$$

if and only if (45) holds. Define

$$r = (I-1)W + IL_{upper}. \qquad (49)$$

From (12)-(14) and (49), it is easy to show that

$$\begin{aligned} r &\geq (I-1)W + \bar{L} \\ &> (I-1)W + \bar{L} - L_1 \end{aligned}$$

and

$$\begin{aligned} r &< (I-1)W + W \\ &= IW \\ &< IW + \bar{L} + 1. \end{aligned}$$

Therefore, r defined in (49) satisfies the inequality in (48) and thus ensures the equation $E_r(\xi) = 0$ in (48). As a result, we can solve $E_r(\xi) = 0$ for ξ by using the MATLAB function *solve*$(\cdot)$, and the obtained ξ will satisfy (45).

However, conducting singular value decomposition on a symbolic matrix requires a large memory space, which may exceed the memory limit of a normal computer. So an alternative approach is needed. It is easy to verify that $\mathbf{T}(\xi)$ is a normal matrix. Also, it is known that the singular values of a normal matrix are equal to the absolute values of its eigenvalues [26]. This means that for a given ξ value, $\mathbf{T}(\xi)$ will have the same number of zero singular values and zero eigenvalues. Based on this observation, we can conduct eigenvalue decomposition, instead of singular value decomposition, on the symbolic matrix $\mathbf{T}(\xi)$, which requires less memory space. Applying the MATLAB function *eig*$(\cdot)$ to $\mathbf{T}(\xi)$ will return an expression *root*$(f(z,\xi), z)$, where $f(z,\xi)$ is a function of order JW with respect to z. This expression implies that the roots of $f(z,\xi)$ with respect to z are the eigenvalues of $\mathbf{T}(\xi)$. As a result, forcing the eigenvalues of $\mathbf{T}(\xi)$ to be zero is equivalent to zeroing the roots of $f(z,\xi)$ with respect to z. Thus, we use the MATLAB function *solve*$(\cdot)$ to solve *root*$(f(z,\xi), z) = 0$ for ξ as follows:

$$solve(root(f(z,\xi), z) == 0, \xi) \qquad (50)$$

which yields JW values. For each obtained ξ value, we substitute it back into (43) and then conduct singular value decomposition for $\mathbf{T}(\xi)$. The ξ value which makes $E_r(\xi) = 0$ is chosen as the desired ξ value. The matrix $\mathbf{T}(\xi)$ constructed based on this desired ξ value is essential to the derivation of the new BSE algorithm.

4) FORMULATION OF ALGORITHM

Since the obtained ξ satisfies (45), it holds $\rho_{11}^{\beta_1}(0) - \xi\cdot\rho_{11}(0) = 0$. Thus, it follows from (47) that

$$\mathbf{T}(\xi) = \boldsymbol{\mathcal{H}} \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & -\xi\cdot\bar{\mathbf{R}}_{\tilde{\mathbf{x}}\tilde{\mathbf{x}}}(0) \end{bmatrix} \boldsymbol{\mathcal{H}}^H. \qquad (51)$$

In this case, the rank of $\mathbf{T}(\xi)$ is $(I-1)W + \bar{L} - L_1 = (I-1)W + \sum_{i=2}^{I} L_i$. Denote the singular vectors corresponding to the K_2 largest singular values of $\mathbf{T}(\xi)$ as $\mathbf{u}_{2,1}, \mathbf{u}_{2,2}, \cdots, \mathbf{u}_{2,K_2}$, where

$$K_2 = (I-1)(W + L_{upper}) \geq (I-1)W + \sum_{i=2}^{I} L_i. \qquad (52)$$

Then we can form the matrix below:

$$\mathbf{U}_2 = \left[\mathbf{u}_{2,1}, \mathbf{u}_{2,2}, \ldots, \mathbf{u}_{2,K_2}\right]. \qquad (53)$$

Based on the matrices $\mathbf{U}_1$ and $\mathbf{U}_2$ shown in (38) and (53), respectively, we define

$$\mathbf{U} = [\mathbf{U}_1, \mathbf{U}_2]. \qquad (54)$$

Assume that $\mathbf{u}$ is the singular vector corresponding to the smallest singular value of $\mathbf{U}$. Then, we propose the following theorem.

Theorem 1: The first element of $\mathbf{u}^H\boldsymbol{\mathcal{H}}$ is nonzero and the rest elements are zero.

Proof: For the $JW \times (IW + \bar{L})$ matrix $\boldsymbol{\mathcal{H}}$ shown in (21), denote the jth column by $\bar{\boldsymbol{h}}_j$ and the (i,j)th entry by $\bar{h}_{i,j}$, where $i = 1, 2, \ldots, JW$ and $j = 1, 2, \ldots, IW + \bar{L}$. The matrix $\mathbf{R}^{\beta_1}_{\tilde{\mathbf{y}}\tilde{\mathbf{y}}}(1)$ in (36) can be further written as

$$
\begin{aligned}
\mathbf{R}^{\beta_1}_{\tilde{\mathbf{y}}\tilde{\mathbf{y}}}(1) &= \rho^{\beta_1}_{11}(0) \cdot \left[\bar{\boldsymbol{h}}_1, \bar{\boldsymbol{h}}_2, \ldots, \bar{\boldsymbol{h}}_{W+L_1}\right] \mathbf{J}_{W+L_1} \begin{bmatrix} \bar{\boldsymbol{h}}_1^H \\ \bar{\boldsymbol{h}}_2^H \\ \vdots \\ \bar{\boldsymbol{h}}_{W+L_1}^H \end{bmatrix} \\
&= \rho^{\beta_1}_{11}(0) \cdot \left[\bar{\boldsymbol{h}}_1, \bar{\boldsymbol{h}}_2, \ldots, \bar{\boldsymbol{h}}_{W+L_1}\right] \begin{bmatrix} 0 \\ \bar{\boldsymbol{h}}_1^H \\ \vdots \\ \bar{\boldsymbol{h}}_{W+L_1-1}^H \end{bmatrix} \\
&= \rho^{\beta_1}_{11}(0) \cdot \left[\bar{\boldsymbol{h}}_2 \bar{\boldsymbol{h}}_1^H + \bar{\boldsymbol{h}}_3 \bar{\boldsymbol{h}}_2^H + \ldots + \bar{\boldsymbol{h}}_{W+L_1} \bar{\boldsymbol{h}}_{W+L_1-1}^H\right] \\
&= \rho^{\beta_1}_{11}(0) \cdot \left[\sum_{i=2}^{W+L_1} \bar{h}^*_{1,i-1} \cdot \bar{\boldsymbol{h}}_i, \ \sum_{i=2}^{W+L_1} \bar{h}^*_{2,i-1} \cdot \bar{\boldsymbol{h}}_i, \ldots, \right. \\
&\qquad \left. \sum_{i=2}^{W+L_1} \bar{h}^*_{JW,i-1} \cdot \bar{\boldsymbol{h}}_i \right]. \qquad (55)
\end{aligned}
$$

From (55), it is easy to see that

$$
\text{span}\left\{\mathbf{R}^{\beta_1}_{\tilde{\mathbf{y}}\tilde{\mathbf{y}}}(1)\right\} = \text{span}\left\{\bar{\boldsymbol{h}}_2, \bar{\boldsymbol{h}}_3, \ldots, \bar{\boldsymbol{h}}_{W+L_1}\right\} \qquad (56)
$$

where $\text{span}\left\{\mathbf{R}^{\beta_1}_{\tilde{\mathbf{y}}\tilde{\mathbf{y}}}(1)\right\}$ denotes the span of the column vectors of $\mathbf{R}^{\beta_1}_{\tilde{\mathbf{y}}\tilde{\mathbf{y}}}(1)$. Similarly, for the matrix $\mathbf{T}(\xi)$ given in (51), we have

$$
\begin{aligned}
&\text{span}\{\mathbf{T}(\xi)\} \\
&\quad = \text{span}\left\{\bar{\boldsymbol{h}}_{W+L_1+1}, \bar{\boldsymbol{h}}_{W+L_1+2}, \ldots, \bar{\boldsymbol{h}}_{IW+\bar{L}}\right\}. \qquad (57)
\end{aligned}
$$

Since

$$
\text{rank}\left\{\mathbf{R}^{\beta_1}_{\tilde{\mathbf{y}}\tilde{\mathbf{y}}}(1)\right\} = W + L_1 - 1
$$

and $\mathbf{u}_{1,1}, \mathbf{u}_{1,2}, \ldots, \mathbf{u}_{1,K_1}$ are the singular vectors corresponding to the K_1 largest singular values of $\mathbf{R}^{\beta_1}_{\tilde{\mathbf{y}}\tilde{\mathbf{y}}}(1)$ with K_1 satisfying (37), it holds that

$$
\text{span}\left\{\mathbf{R}^{\beta_1}_{\tilde{\mathbf{y}}\tilde{\mathbf{y}}}(1)\right\} \subseteq \text{span}\left\{\mathbf{u}_{1,1}, \mathbf{u}_{1,2}, \ldots, \mathbf{u}_{1,K_1}\right\}. \qquad (58)
$$

Moreover, since

$$
\text{rank}\{\mathbf{T}(\xi)\} = (I-1)W + \sum_{i=1}^{I-1} L_i
$$

and $\mathbf{u}_{2,1}, \mathbf{u}_{2,2}, \ldots, \mathbf{u}_{2,K_2}$ are the singular vectors corresponding to the K_2 largest singular values of $\mathbf{T}(\xi)$ with K_2 satisfying (52), we have

$$
\text{span}\{\mathbf{T}(\xi)\} \subseteq \text{span}\left\{\mathbf{u}_{2,1}, \mathbf{u}_{2,2}, \ldots, \mathbf{u}_{2,K_2}\right\}. \qquad (59)
$$

It is easy to find that the matrix $\mathbf{U}$ defined in (54) has JW rows and $K_1 + K_2$ columns. Since $\boldsymbol{\mathcal{H}}$ is a tall matrix, it holds that

$$
K_1 + K_2 \leq JW - 1.
$$

Hence, the singular vector $\mathbf{u}$ corresponding to the smallest singular value of $\mathbf{U}$ must satisfy

$$
\mathbf{u}^H \mathbf{U} = \mathbf{0}.
$$

From (56)-(59), it follows that

$$
\mathbf{u}^H \bar{\boldsymbol{h}}_k = 0, \ k = 2, 3, \ldots, IW + \bar{L}. \qquad (60)
$$

Since $\boldsymbol{\mathcal{H}}$ is of full column rank, its first column $\bar{\boldsymbol{h}}_1$ is independent of all other columns $\bar{\boldsymbol{h}}_2, \bar{\boldsymbol{h}}_3, \ldots, \bar{\boldsymbol{h}}_{IW+\bar{L}}$. Therefore, it holds with probability one that

$$
\mathbf{u}^H \bar{\boldsymbol{h}}_1 \neq 0. \qquad (61)
$$

Combining (60) and (61) completes the proof. □

From Theorem 1, the vector $\mathbf{u}$ is not orthogonal to the first column of $\boldsymbol{\mathcal{H}}$ but orthogonal to the other columns of $\boldsymbol{\mathcal{H}}$. Thus, $\mathbf{u}$ is a desired extraction vector that can recover the SOI, $x_1(n)$, from $\mathbf{y}(n)$.

In summary, the proposed BSE algorithm is formulated as follows:

- Step 1: Choose a slide-window width W satisfying (14), followed by constructing $\tilde{\mathbf{y}}_1(n), \tilde{\mathbf{y}}_2(n), \cdots, \tilde{\mathbf{y}}_J(n)$ by (11) and $\tilde{\mathbf{y}}(n)$ by (17).
- Step 2: Compute the cyclic correlation matrices $\mathbf{R}^{\beta_1}_{\tilde{\mathbf{y}}\tilde{\mathbf{y}}}(0)$ and $\mathbf{R}^{\beta_1}_{\tilde{\mathbf{y}}\tilde{\mathbf{y}}}(1)$ by (35), and the autocorrelation matrix $\mathbf{R}_{\tilde{\mathbf{y}}\tilde{\mathbf{y}}}(0)$ by $\mathbf{R}_{\tilde{\mathbf{y}}\tilde{\mathbf{y}}}(0) = \langle \tilde{\mathbf{y}}(n)\tilde{\mathbf{y}}^H(n)\rangle$.
- Step 3: Estimate the noise variance σ_w^2, which is the smallest eigenvalue of $\mathbf{R}_{\tilde{\mathbf{y}}\tilde{\mathbf{y}}}(0)$, and then remove noise component from $\mathbf{R}_{\tilde{\mathbf{y}}\tilde{\mathbf{y}}}(0)$ to obtain $\bar{\mathbf{R}}_{\tilde{\mathbf{y}}\tilde{\mathbf{y}}}(0)$ by (41).
- Step 4: Based on $\mathbf{R}^{\beta_1}_{\tilde{\mathbf{y}}\tilde{\mathbf{y}}}(0)$ and $\bar{\mathbf{R}}_{\tilde{\mathbf{y}}\tilde{\mathbf{y}}}(0)$, conduct a series of calculations: i) form the symbolic matrix $\mathbf{T}(\xi)$ with the variable ξ by (43) and compute r by (49); ii) apply $eig(\cdot)$ to the symbolic matrix $\mathbf{T}(\xi)$ to get the expression $root(f(z,\xi), z)$; iii) use (50) to obtain a set of ξ values; iv) substitute each ξ value back into (43), conduct singular value decomposition for $\mathbf{T}(\xi)$, and select the ξ value making $E_r(\xi) = 0$ as the desired ξ value; v) apply the desired ξ value to (43) to calculate $\mathbf{T}(\xi)$.
- Step 5: Find $\mathbf{u}_{1,1}, \mathbf{u}_{1,2}, \ldots, \mathbf{u}_{1,K_1}$, which are the singular vectors corresponding to the K_1 largest singular values of $\mathbf{R}^{\beta_1}_{\tilde{\mathbf{y}}\tilde{\mathbf{y}}}(1)$. Then, form the matrix $\mathbf{U}_1$ by (38).
- Step 6: Find $\mathbf{u}_{2,1}, \mathbf{u}_{2,2}, \ldots, \mathbf{u}_{2,K_2}$, which are the singular vectors corresponding to the K_2 largest singular values of $\mathbf{T}(\xi)$. Then, form the matrix $\mathbf{U}_2$ by (53).
- Step 7: Construct the matrix $\mathbf{U}$ by (54). The singular vector corresponding to the smallest singular value of $\mathbf{U}$ is chosen as the extraction vector $\mathbf{u}$.

Remark: The proposed BSE algorithm does not require the exact values of $L_1, L_2, \cdots, L_I$ but their upper bound L_{upper}. So, in the implementation of the algorithm, we can assume, without loss generality, that all of the columns of $\mathbf{H}(z)$ are of equal order L_{upper}.

IV. SIMULATION RESULTS

In this section, we present simulation examples to illustrate the performance of the proposed SOCS-based BSE method, in comparison with the methods in [6], [18] and [24]. In the simulations, the source signals are generated as follows:

$$x_1(n) = s_1(n) * \cos(\alpha_1 \cdot n)$$
$$x_i(n) = s_i(n), \quad i = 2, 3, \ldots, I$$

where $s_1(n), s_2(n), \ldots, s_I(n)$ are randomly generated temporarily white sequences. Thus the SOI, $x_1(n)$, is cyclostationary with cyclic frequency $\beta_1 = 2\alpha_1$ and $\rho_{11}(0) = 2\rho_{11}^{\beta_1}(0)$. The MIMO FIR mixing matrix $\mathbf{H}(z)$ is randomly generated in each simulation run. Randomly generated white Gaussian noise is added to the source signal mixtures and the signal-to-noise ratio (SNR) is defined as

$$\text{SNR} = -10log_{10}\left(\sigma_w^2\right).$$

Ideally, for the compound vector $\mathbf{c}$ defined in (22), only one of its first W elements is nonzero and the rest elements should be zero. However, this is not practically achievable due to finite sample size and computational inaccuracy. Therefore, the performance of our method is measured by means of the mean interference rejection level (MIRL) of the extraction vector. Let

$$\mathbf{c}^T = [c_1, c_2, \ldots, c_k, \ldots, c_{IW+\bar{L}}]$$

and assume c_k is the maximum element out of the first W elements of $\mathbf{c}$. The MIRL index is defined as follows:

$$\text{MIRL(dB)} = 20\log_{10}\left(\frac{1}{(IW+\bar{L}-1)\cdot|c_k|}\sum_{i=1,i\neq k}^{IW+\bar{L}}|c_i|\right).$$

We compute MIRL by averaging 200 independent runs. Clearly, the smaller MIRL, the better SOI extraction performance.

In the first simulation, we consider an MIMO FIR mixing system of 3 inputs, 5 outputs and order 2, i.e., $I = 3$, $J = 5$ and $L = 2$. The SNR is kept at SNR= 25dB. Fig. 1 shows the MIRL of the proposed BSE method versus sample size N. It can be seen that satisfactory performance is achieved for all tested sample sizes. Also, with the increase of sample size, the MIRL decreases accordingly, as expected. Our method exploits the second-order statistics properties of the system outputs, specifically the output cyclic-correlation and autocorrelation matrices. So using more data samples will generate more accurate statistics properties and thus produce better source extraction performance.

The second simulation evaluates the performance of the proposed method versus the order of the mixing system, where $I = 3$, $J = 5$, $N = 30000$ and SNR= 25dB. Fig. 2 shows the simulation result. We can see that the MIRL increases with the rise of the system order. This is not surprising as higher system order means that the system is more complicated, which will lower the performance of BSE. Nevertheless, our method performs well even in the cases of relatively high system orders, e.g., the proposed method achieves an MIRL of about −15dB when the order of the system is 7.

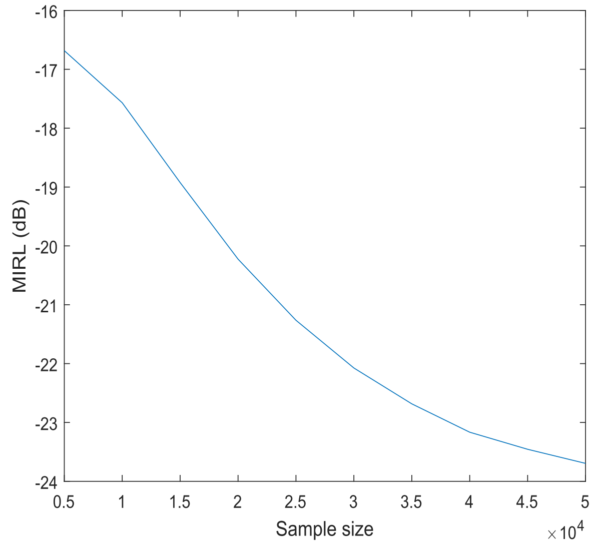

FIGURE 1. The MIRL of the proposed method versus sample size, where $I = 3$, $J = 5$, $L = 2$ and SNR= 25dB.

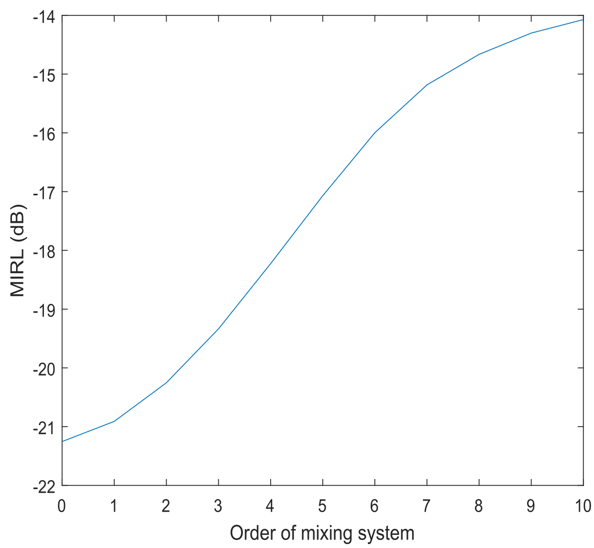

FIGURE 2. The MIRL of the proposed method versus system order, where $I = 3$, $J = 5$, $N = 30000$ and SNR= 25dB.

Then, the MIRL of our method is assessed under different number of sources (or inputs), where the other simulation parameters are $J = 7$, $L = 2$, $N = 30000$ and SNR= 25dB. As shown in Fig. 3, on the one hand, increasing the number of sources enlarges MIRL, degrading the accuracy of SOI extraction. On the other hand, when the number of sources is smaller than the number of outputs, which is 7 in this simulation, the MIRL is satisfactory. However, when the source number reaches 7, the extraction performance is quite poor. This is understandable as the assumption A2) is not satisfied in this case.

Finally, we compare the proposed method with those methods in [6], [18] and [24]. Fig. 4 shows the MIRLs of the four methods versus SNR, where $I = 3$, $J = 5$, $L = 2$ and $N = 30000$. It can be seen that our method performs

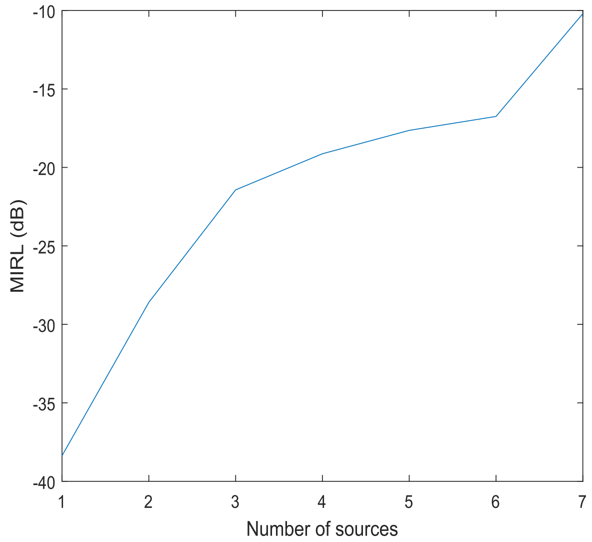

FIGURE 3. The MIRL of the proposed method versus the number of sources, where $J = 7$, $L = 2$, $N = 30000$ and SNR= 25dB.

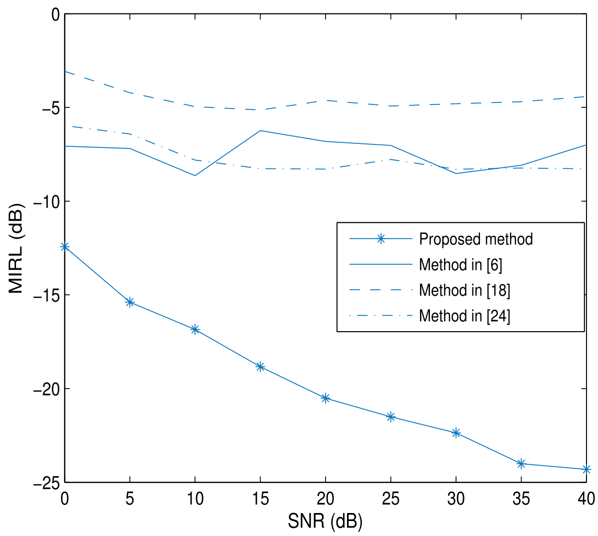

FIGURE 4. The MIRLs of four compared methods versus SNR, where $I = 3$, $J = 5$, $L = 2$ and $N = 30000$.

very well at moderate and high SNR levels. It also achieves satisfactory performance when the SNR is not too low, thanks to the noise removal procedure actioned in (41). For example, it achieves an MIRL of about -15.5dB at SNR=5dB. In contrast, the other three methods fail to demonstrate a good performance. The main reasons are as follows. The method in [6] is a higher-order statistics based method and it requires the sources to have constant modulus. The method in [18] conducts signal separation in distinct frequency bins of the frequency domain, which introduces a frequency-dependent permutation ambiguity. Correcting this permutation ambiguity requires the sources to possess some special properties. However, the condition required by [6] and some of the conditions required by [18] do not exist in the source signals used in our simulation. Furthermore, the ambiguity correction procedure in [18] itself also introduces additional errors to the final outcome of BSE. Regarding the method in [24], it assumes that the entries of the first column of the mixing matrix are constants instead of polynomials. However, the mixing matrix used in the simulation is more general and all of its entries are polynomials.

V. CONCLUSION

This paper deals with the problem of extracting an SOI that is cyclostationary from the outputs of an unknown MIMO FIR mixing system. Although great efforts have been made to tackle this problem and some methods have been reported in the literature [15]–[18], [22]–[24], they have various disadvantages. Specifically, they rely on the HOCS of the measured data [15]–[17], [22] which is costly in computation, have local minimum problem [15]–[17], suffer from the frequency-dependent permutation ambiguity problem [18], struggle with the non-removable noise contamination problem [23], or impose restrictive conditions on the MIMO FIR mixing system (i.e., the entries of the first column of the mixing matrix are constants) and the SOI (i.e., the ratio of the SOI's cyclic correlation function over its correlation function is known) [24]. In this paper, a new SOCS-based method is proposed to recover the SOI from MIMO FIR mixtures. The proposed method overcomes the problems existing in the methods in [15]–[18] and [22]–[24]. Simulation results show the superior performance of the new method over the existing methods.

REFERENCES

[1] 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.
[2] N. Bouguerriou, M. Haritopoulos, C. Capdessus, and L. Allam, "Novel cyclostationarity-based blind source separation algorithm using second order statistical properties: Theory and application to the bearing defect diagnosis," *Mech. Syst. Signal Process.*, vol. 19, no. 6, pp. 1260–1281, Nov. 2005.
[3] Y. Xiang, "Blind source separation based on constant modulus criterion and signal mutual information," *Comput. Elect. Eng.*, vol. 34, no. 5, pp. 416–422, Sep. 2008.
[4] Y. Xiang, D. Peng, Y. Xiang, and S. Guo, "Novel Z-domain precoding method for blind separation of spatially correlated signals," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 24, no. 1, pp. 94–105, Jan. 2013.
[5] K. Abed-Meraim, Y. Xiang, J. H. Manton, and Y. Hua, "Blind source-separation using second-order cyclostationary statistics," *IEEE Trans. Signal Process.*, vol. 49, no. 4, pp. 694–701, Apr. 2001.
[6] E. Florian, A. Chevreuil, and P. Loubaton, "Blind source separation of convolutive mixtures of non-circular linearly modulated signals with unknown baud rates," *Signal Process.*, vol. 92, no. 3, pp. 715–726, Mar. 2011.
[7] W. Gardner, "An introduction to cyclostationary signals," in *Cyclostationarity in Communications and Signal Processing*, W. A. Gardner, Ed. Piscataway, NJ, USA: IEEE Press, 1994, ch. 1, pp. 1–90.
[8] Y. Xiang, W. Yu, J. Zhang, and S. An, "Blind source separation based on phase and frequency redundancy of cyclostationary signals," *IEICE Trans. Fundam. Electron., Commun. Comput. Sci.*, vol. E87-A, no. 12, pp. 3343–3349, Dec. 2004.
[9] N. CheViet, M. El Badaoui, A. Belouchrani, and F. Guillet, "Blind separation of cyclostationary sources using non-orthogonal approximate joint diagonalization," in *Proc. 5th IEEE Sensor Array Multichannel Signal Process. Workshop*, Darmstadt, Germany, Jul. 2008, pp. 492–495.
[10] A. Ferréol, P. Chevalier, and L. Albera, "Second-order blind separation of first- and second-order cyclostationary sources-application to AM, FSK, CPFSK, and deterministic sources," *IEEE Trans. Signal Process.*, vol. 52, no. 4, pp. 845–861, Apr. 2004.

[11] G. Jie and G. Yan, "A fast convergence blind source separation algorithm based on cyclostationary statistics," in *Proc. 12th IEEE Int. Conf. Commun. Technol.*, Nanjing, China, Nov. 2010, pp. 1472–1475.

[12] S. Rhioui, N. Thirion-Moreau, and E. Moreau, "Under-determined blind identification of cyclo-stationary signals with unknown cyclic frequencies," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process.*, vol. 4. Toulouse, France, May 2006, pp. 553–556.

[13] R. Wang, C. Hou, and J. Yang, "Blind adaptive beamforming algorithm based on cyclostationary signals," in *Proc. Int. Conf. Image Anal. Signal Process.*, Zhejiang, China, 2010, pp. 488–491.

[14] R. Wang, C. Hou, and D. Chen, "Blind separation of instantaneous linear mixtures of cyclostationary signals," in *Proc. Int. Conf. Image Anal. Signal Process.*, Zhejiang, China, 2010, pp. 492–495.

[15] H. Ghennioui, N. Thirion-Moreau, E. Moreau, D. Aboutajdine, and A. Adib, "A novel approach based on non-unitary joint block-diagonalization for the blind MIMO equalization of cyclo-stationary signals," in *Proc. 16th Eur. Signal Process. Conf.*, Lausanne, Switzerland, 2008, pp. 1–5.

[16] K. Wei and Z. Yongxiang, "A new algorithm for cyclostationary signal blind deconvolution," in *Proc. Int. Conf. Consum. Electron., Commun. Netw.*, Xianning, China, 2011, pp. 3835–3838.

[17] W. Wang, M. G. Jafari, S. Sanei, and J. A. Chambers, "Blind separation of convolutive mixtures of cyclostationary sources using an extended natural gradient method," in *Proc. 7th Int. Symp. Signal Process. Appl.*, Paris, France, 2003, pp. 93–96.

[18] K. Sabri, M. El Badaoui, F. Guillet, A. Adib, and D. Aboutajdine, "A frequency domain-based approach for blind MIMO system identification using second-order cyclic statistics," *Signal Process.*, vol. 89, no. 1, pp. 77–86, Jan. 2009.

[19] F. Ghaderi, K. Nazarpour, J. G. McWhirter, and S. Sanei, "Removal of ballistocardiogram artifacts using the cyclostationary source extraction method," *IEEE Trans. Biomed. Eng.*, vol. 57, no. 11, pp. 2667–2676, Nov. 2010.

[20] F. Ghaderi, B. Makkiabadi, J. G. McWhirter, and S. Sanei, "Blind source extraction of cyclostationary sources with common cyclic frequencies," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process.*, Dallas, TX, USA, Mar. 2010, pp. 4146–4149.

[21] F. Ghaderi, H. R. Mohseni, J. G. McWhirter, and S. Sanei, "Blind source extraction of periodic signals," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process.*, Taipei, Taiwan, Apr. 2009, pp. 377–380.

[22] L. Castedo and A. R. Figueiras-Vidal, "An adaptive beamforming technique based on cyclostationary signal properties," *IEEE Trans. Signal Process.*, vol. 43, no. 7, pp. 1637–1650, Jul. 1995.

[23] R. Boustany and J. Antoni, "A subspace method for the blind extraction of a cyclostationary source: Application to rolling element bearing diagnostics," *Mech. Syst. Signal Process.*, vol. 19, no. 6, pp. 1245–1259, Nov. 2005.

[24] Y. Xiang, I. Ubhayaratne, Z. Yang, B. Rolfe, and D. Peng, "Blind extraction of cyclostationary signal from convolutional mixtures," in *Proc. 9th IEEE Conf. Ind. Electron. Appl.*, Hangzhou, China, Jun. 2014, pp. 857–861.

[25] R. Bitmead, S.-Y. Kung, B. Anderson, and T. Kailath, "Greatest common divisor via generalized Sylvester and Bezout matrices," *IEEE Trans. Autom. Control*, vol. 23, no. 6, pp. 1043–1047, Dec. 1978.

[26] A. Ruhe, "On the closeness of eigenvalues and singular values for almost normal matrices," *Linear Algebra Appl.*, vol. 11, no. 1, pp. 87–93, 1975.

YONG XIANG (SM'12) received the Ph.D. degree in electrical and electronic engineering from The University of Melbourne, Australia. He is currently a Professor and the Director of the Artificial Intelligence and Data Analytics Research Cluster, School of Information Technology, Deakin University, Australia. His research interests include information security and privacy, multimedia (speech/image/video) processing, wireless sensor networks, massive MIMO, and biomedical signal processing. He has authored over 80 refereed journal articles and numerous conference papers in these areas. He has served as the Program Chair, the TPC Chair, the Symposium Chair, and the Session Chair of a number of international conferences. He is an Associate Editor of the IEEE Signal Processing Letters and the IEEE Access.

DEZHONG PENG (M'09) received the B.S. degree in applied mathematics and the M.S. and Ph.D. degrees in computer software and theory from the University of Electronic Science and Technology of China, China, in 1998, 2001, and 2006, respectively. From 2001 to 2007, he was with the University of Electronic Science and Technology of China, as an Assistant Lecturer and a Lecturer. He was a Post-Doctoral Research Fellow with the School of Engineering, Deakin University, Australia, from 2007 to 2009. He is currently a Professor with the Machine Intelligence Laboratory, College of Computer Science, Sichuan University, Chengdu, China. His research interests include blind signal processing and neural networks.

INDIVARIE UBHAYARATNE received the B.Sc. degree (Hons.) in engineering from the University of Moratuwa, Sri Lanka, in 2011. She is currently pursuing the Ph.D. degree at Deakin University, Australia. Her research interests include blind signal processing, system modeling, and software engineering.

BERNARD ROLFE received the Economics and Engineering degree (Hons.) and the Ph.D. degree from The Australian National University in 1995 and 2002, respectively. He is currently an Associate Professor with Deakin University, Australia. He is also the Lightweighting Team Leader of Excellerate Australia (Automotive Technology Cooperative Research Centre). He has authored over 150 refereed articles. His research focus is the forming of lightweight structures, including the development of innovative forming methods.

MICHAEL PEREIRA received the B.Eng. degree (Hons.) from RMIT University, Australia, in 2002, and the Ph.D. degree from Deakin University, Australia, in 2010. He was a Research Fellow with Deakin University from 2010 to 2014, where he is currently a Lecturer. He has authored over 25 refereed articles. His research interests include tool wear analysis and metal stamping modeling.