CIRCUITS SYSTEMS SIGNAL PROCESSING
2007

DOI: 10.1007/s00034-007-4003-y

ROBUST EIGENVECTOR ALGORITHMS FOR BLIND DECONVOLUTION OF MIMO LINEAR SYSTEMS*

Mitsuru Kawamoto,[1,2] *Kiyotaka Kohno,*[3] *and Yujiro Inouye*[3,4]

Abstract. This paper presents eigenvector algorithms (EVAs) for blind deconvolution of multiple-input, multiple-output infinite impulse response channels (convolutive mixtures). An attractive feature of one of the proposed EVAs is that it is insensitive to Gaussian noises which are added to the outputs of the channels; hence, the proposed EVA is referred to as a "robust" eigenvector algorithm. Simulation results show the validity of the proposed EVAs.
Key words: Eigenvector algorithms, robust eigenvector algorithms, blind deconvolution, multiple-input, multiple-output infinite impulse response channels, Gaussian noise.

1. Introduction

In the last decade, blind signal processing (BSP), which includes, e.g., blind source separation (BSS), blind deconvolution (BD), and blind equalization (BE), has attracted many researchers in the fields of speech signal processing, communication systems, biomedical signal processing, and so on. One of the main reasons may be that a BSP technique can extract a set of source signals from observations

* Received July 19, 2006; revised January 29, 2007; This work was supported by Grant-in-Aids for Scientific Research by the Ministry of Education, Culture, Sports, Science, and Technology of Japan, no. 18500146 and no. 18500054.

[1] National Institute of Advanced Industrial Science and Technology (AIST), Central 2, 1-1-1 Umezono, Tsukuba, Ibaraki 305-8568, Japan.

[2] Laboratory for Biologically Integrative Sensors, Bio-Mimetic Control Research Center, RIKEN, 2271-130, Anagahora, Shimoshidami, Moriyama-ku, Nagoya, 463-0003, Japan. E-mail: m.kawamoto@aist.go.jp (corresponding author)

[3] Department of Electronic and Control Systems Engineering, Shimane University, 1060 Nishikawatsu, Matsue, Shimane 690-8504, Japan. E-mail for Kohno: kohno@yonago-k.ac.jp; E-mail for Inouye: inouye@riko.shimane-u.ac.jp

[4] Laboratory for Advanced Brain Signal Processing, Brain Science Institute, RIKEN, 2-1 Hirosawa, Wako, Saitama 351-0198, Japan.

that are mixtures of the source signals, even if neither the particular information on source signals nor the mixing processes are known a priori.

In this paper, we deal with the BD problem for multiple-input, multiple-output (MIMO) infinite impulse response (IIR) channels. To solve this problem, we use eigenvector algorithms (EVAs) [6], [7], [15]. The first proposal of an EVA was done by Jelonnek et al. [6]. They proposed the EVA for solving BE problems of single-input, single-output (SISO) channels or single-input, multiple-output (SIMO) channels. In [15], several procedures for the BSS of instantaneous mixtures, using the generalized eigenvalue decomposition, have been introduced. Recently, the authors have proposed an EVA that can solve BSS problems in the case of MIMO static systems (instantaneous mixtures) [9], [10].

In this paper, based on the idea in [9], [10], we shall show that EVAs can be used to solve the BD problem of MIMO-IIR channels, where the proposed EVAs have some technical difficulties that are different from our conventional EVAs in [9], [10] (see Section 3). Moreover, it will be shown that the proposed EVA can be modified to treat noisy situations such that the BD can be achieved with as little influence of Gaussian noise as possible; hence this type of EVA is referred to as a "robust" EVA (REVA). Computer simulations are presented to demonstrate the validity of the proposed EVA and REVA.

This paper uses the following notation. Let Z denote the set of all integers. Let C denote the set of all complex numbers. Let $\boldsymbol{C}^n$ denote the set of all n-column vectors with complex components. Let $\boldsymbol{C}^{m\times n}$ denote the set of all $m \times n$ matrices with complex components. The superscripts T, $*$, H, and $\dagger$ denote, respectively, the transpose, the complex conjugate, the complex conjugate transpose (Hermitian) and the (Moore-Penrose) pseudoinverse operations of a matrix. The symbol $\boldsymbol{I}$ denotes an identity matrix. The symbols block-diag$\{\cdots\}$ and diag$\{\cdots\}$ denote respectively a block diagonal and a diagonal matrix with the block diagonal and the diagonal elements $\{\cdots\}$. The symbol cum$\{x_1,x_2,x_3,x_4\}$ denotes a fourth-order cumulant of x_i's. Let $i = \overline{1, n}$ stand for $i = 1, 2, \ldots, n$.

2. Problem formulation and assumptions

We consider a MIMO channel with n inputs and m outputs as described by

$$\boldsymbol{y}(t) = \sum_{k=-\infty}^{\infty} \boldsymbol{H}^{(k)}\boldsymbol{s}(t-k) + \boldsymbol{n}(t), \qquad t \in Z, \tag{1}$$

where $\boldsymbol{s}(t)$ is an n-column vector of input (or source) signals, $\boldsymbol{y}(t)$ is an m-column vector of channel outputs, $\boldsymbol{n}(t)$ is an m-column vector of Gaussian noises, and $\{\boldsymbol{H}^{(k)}\}$ is an $m \times n$ impulse response matrix sequence. The transfer function of the

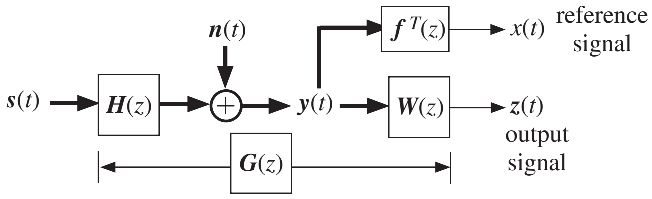

Figure 1. The composite system of an unknown system and a deconvolver, and a reference filter.

channel is defined by

$$\boldsymbol{H}(z) = \sum_{k=-\infty}^{\infty} \boldsymbol{H}^{(k)} z^k, \qquad z \in C. \tag{2}$$

To recover the source signals, we process the output signals by an $n \times m$ deconvolver (or equalizer) $\boldsymbol{W}(z)$ described by

$$\begin{aligned} \boldsymbol{z}(t) &= \sum_{k=-\infty}^{\infty} \boldsymbol{W}^{(k)} \boldsymbol{y}(t-k) \\ &= \sum_{k=-\infty}^{\infty} \boldsymbol{G}^{(k)} \boldsymbol{s}(t-k) + \sum_{k=-\infty}^{\infty} \boldsymbol{W}^{(k)} \boldsymbol{n}(t-k), \end{aligned} \tag{3}$$

where $\{\boldsymbol{G}^{(k)}\}$ is the impulse response matrix sequence of $\boldsymbol{G}(z) := \boldsymbol{W}(z)\boldsymbol{H}(z)$, which is defined by

$$\boldsymbol{G}(z) = \sum_{k=-\infty}^{\infty} \boldsymbol{G}^{(k)} z^k, \qquad z \in C. \tag{4}$$

The cascade connection of the unknown system and the deconvolver is illustrated in Figure 1.

The objective of multichannel blind deconvolution is to construct a deconvolver that recovers the original source signals only from their convolutive mixtures.

We put the following assumptions on the channel, the source signals, the deconvolver, and the noises.

(A1) The transfer function $\boldsymbol{H}(z)$ is stable and has full column rank on the unit circle $|z| = 1$, where the assumption **A1** implies that the unknown system has less inputs than outputs, i.e., $n \leq m$, and there exists a left stable inverse of the unknown system.

(A2) The input sequence $\{\boldsymbol{s}(t)\}$ is a complex, zero-mean, and non-Gaussian random vector process with element processes $\{s_i(t)\}$, $i = \overline{1, n}$ being mutually independent. Each element process $\{s_i(t)\}$ is an i.i.d. process with a variance $\sigma_{s_i}^2 \neq 0$ and a nonzero fourth-order cumulant $\gamma_i \neq 0$ defined as

$$\gamma_i = \text{cum}\{s_i(t), s_i(t), s_i^*(t), s_i^*(t)\} \neq 0. \tag{5}$$

(A3) The deconvolver $\boldsymbol{W}(z)$ is a finite impulse response (FIR) channel of sufficient length L so that the truncation effect can be ignored.

(A4) The noise sequence $\{\boldsymbol{n}(t)\}$ is a zero-mean, Gaussian vector stationary process whose component processes $\{n_j(t)\}$, $j = \overline{1, m}$ have nonzero variances $\sigma_{n_j}^2$, $j = \overline{1, m}$.

(A5) The two vector sequences $\{\boldsymbol{n}(t)\}$ and $\{\boldsymbol{s}(t)\}$ are mutually statistically independent.

Under **A3**, the impulse responses $\boldsymbol{G}^{(k)}$ for $k \in Z$ of the cascade system are given by

$$\boldsymbol{G}^{(k)} := \sum_{\tau=L_1}^{L_2} \boldsymbol{W}^{(\tau)} \boldsymbol{H}^{(k-\tau)}, \qquad k \in Z, \tag{6}$$

where the length $L := L_2 - L_1 + 1$ is taken to be sufficiently large. In a vector form, (6) can be written as

$$\tilde{\boldsymbol{g}}_i = \tilde{\boldsymbol{H}} \tilde{\boldsymbol{w}}_i, \qquad i = \overline{1, n}, \tag{7}$$

where $\tilde{\boldsymbol{g}}_i$ is the column vector consisting of the ith output impulse response of the cascade system defined by

$$\tilde{\boldsymbol{g}}_i := \left[\boldsymbol{g}_{i1}^T, \boldsymbol{g}_{i2}^T, \ldots, \boldsymbol{g}_{in}^T\right]^T, \tag{8}$$

$$\boldsymbol{g}_{ij} := \left[\ldots, g_{ij}(-1), g_{ij}(0), g_{ij}(1), \ldots\right]^T, \qquad j = \overline{1, n}, \tag{9}$$

where $g_{ij}(k)$ is the (i, j)th element of matrix $\boldsymbol{G}^{(k)}$, and $\tilde{\boldsymbol{w}}_i$ is the mL-column vector consisting of the tap coefficients (corresponding to the ith output) of the deconvolver defined by

$$\tilde{\boldsymbol{w}}_i := \left[\boldsymbol{w}_{i1}^T, \boldsymbol{w}_{i2}^T, \ldots, \boldsymbol{w}_{im}^T\right]^T \in \boldsymbol{C}^{mL}, \tag{10}$$

$$\boldsymbol{w}_{ij} := \left[w_{ij}(L_1), w_{ij}(L_1 + 1), \ldots, w_{ij}(L_2)\right]^T \in \boldsymbol{C}^L, \qquad j = \overline{1, m}, \tag{11}$$

where $w_{ij}(k)$ is the (i, j)th element of matrix $\boldsymbol{W}^{(k)}$, and $\tilde{\boldsymbol{H}}$ is the $n \times m$ block matrix defined by

$$\tilde{\boldsymbol{H}} := \begin{bmatrix} \boldsymbol{H}_{11} & \boldsymbol{H}_{12} & \cdots & \boldsymbol{H}_{1m} \\ \boldsymbol{H}_{21} & \boldsymbol{H}_{22} & \cdots & \boldsymbol{H}_{2m} \\ \vdots & \vdots & \vdots & \vdots \\ \boldsymbol{H}_{n1} & \boldsymbol{H}_{n2} & \cdots & \boldsymbol{H}_{nm} \end{bmatrix} \tag{12}$$

whose (i, j)th block element $\boldsymbol{H}_{ij}$ is the matrix (of L columns and possibly infinite number of rows) with the (l, r)th element $[\boldsymbol{H}_{ij}]_{lr}$ defined by

$$\begin{aligned} &[\boldsymbol{H}_{ij}]_{lr} := h_{ji}(l - r), \\ &l = 0, \pm 1, \pm 2, \ldots, \qquad r = \overline{L_1, L_2}, \end{aligned} \tag{13}$$

where $h_{ij}(k)$ is the (i, j)th element of the matrix $\boldsymbol{H}^{(k)}$.

In the multichannel blind deconvolution problem, we want to adjust $\tilde{\boldsymbol{w}}_i$'s ($i = \overline{1, n}$) so that

$$[\tilde{\boldsymbol{g}}_1, \ldots, \tilde{\boldsymbol{g}}_n] = \tilde{\boldsymbol{H}}[\tilde{\boldsymbol{w}}_1, \ldots, \tilde{\boldsymbol{w}}_n] = [\tilde{\boldsymbol{\delta}}_1, \ldots, \tilde{\boldsymbol{\delta}}_n]\boldsymbol{P}, \tag{14}$$

where $\boldsymbol{P}$ is an $n \times n$ permutation matrix, and $\tilde{\boldsymbol{\delta}}_i$ is the n-block column vector defined by

$$\tilde{\boldsymbol{\delta}}_i := [\boldsymbol{\delta}_{i1}^T, \boldsymbol{\delta}_{i2}^T, \ldots, \boldsymbol{\delta}_{in}^T]^T, \qquad i = \overline{1, n} \tag{15}$$

$$\boldsymbol{\delta}_{ij} := \begin{cases} \hat{\boldsymbol{\delta}}_i, & \text{if} \quad i = j, \\ (\cdots, 0, 0, 0, \cdots)^T, & \text{otherwise.} \end{cases} \tag{16}$$

Here, $\hat{\boldsymbol{\delta}}_i$ is the column vector (of infinite elements) whose rth element $\hat{\delta}_i(r)$ is given by

$$\hat{\delta}_i(r) = d_i \delta(r - k_i), \tag{17}$$

where $\delta(t)$ is the Kronecker delta function, d_i is a complex number standing for a scale change and a phase shift, and k_i is an integer standing for a time shift.

3. Eigenvector algorithms (EVAs)

3.1. *Analysis of EVAs with reference signals for MIMO-IIR channels*

In this subsection, we assume that there is no noise $\boldsymbol{n}(t)$ in the output $\boldsymbol{y}(t)$, and then analyze EVAs for the MIMO channel (1). Under this assumption, to solve the BD problem, the following cross-cumulant between $z_i(t)$ and a reference signal $x(t)$ (see Figure 1) is defined:

$$D_{z_i x} = \mathrm{cum}\{z_i(t), z_i^*(t), x(t), x^*(t)\}, \tag{18}$$

where $z_i(t)$ is the ith element of $\boldsymbol{z}(t)$ in (3) and the reference signal $x(t)$ is given by $\boldsymbol{f}^T(z)\boldsymbol{y}(t) := \sum_{j=1}^{m} \sum_{k=L_1}^{L_2} f_j(k) y_j(t-k)$, using an appropriate filter $\boldsymbol{f}(z)$, where $\boldsymbol{f}(z)$ is an m-column vector whose elements are $f_j(z) = \sum_{k=L_1}^{L_2} f_j(k) z^k$, $j = \overline{1, m}$. The filter $\boldsymbol{f}(z)$ is called a *reference filter*. Let $\boldsymbol{a}(z) := \boldsymbol{H}^T(z)\boldsymbol{f}(z) = [a_1(z), a_2(z), \ldots, a_n(z)]^T$, then $x(t) = \boldsymbol{f}^T(z)\boldsymbol{H}(z)\boldsymbol{s}(t) = \boldsymbol{a}^T(z)\boldsymbol{s}(t)$. The element $a_i(z)$ of the filter $\boldsymbol{a}(z)$ is defined as $a_i(z) = \sum_{k=-\infty}^{\infty} a_i(k) z^k$. Then it can be seen from (7) that the equation $\boldsymbol{a}(z) = \boldsymbol{H}^T(z)\boldsymbol{f}(z)$ can be rewritten in vector notation as

$$\tilde{\boldsymbol{a}} = \tilde{\boldsymbol{H}}\tilde{\boldsymbol{f}}, \tag{19}$$

where $\tilde{\boldsymbol{a}}$ is the vector whose elements are the parameters $a_i(k)$ of the filter $a_i(z)$, that is,

$$\tilde{\boldsymbol{a}} := \left[\boldsymbol{a}_1^T, \boldsymbol{a}_2^T, \ldots, \boldsymbol{a}_n^T\right]^T, \tag{20}$$

$$\boldsymbol{a}_i := [\ldots, a_i(-1), a_i(0), a_i(1), \ldots]^T, \tag{21}$$

$\tilde{\boldsymbol{f}}$ is the mL-column vector whose elements are the parameters $f_j(k)$ of $f_j(z)$, defined by

$$\tilde{\boldsymbol{f}} := \left[\boldsymbol{f}_1^T, \boldsymbol{f}_2^T, \ldots, \boldsymbol{f}_m^T\right]^T \in \boldsymbol{C}^{mL}, \tag{22}$$

$$\boldsymbol{f}_j := \left[f_j(L_1), f_j(L_1+1), \ldots, f_j(L_2)\right]^T \in \boldsymbol{C}^L, \qquad j = \overline{1, m}, \tag{23}$$

and $\tilde{\boldsymbol{H}}$ in (19) is the same $n \times m$ block matrix as (12).

Remark 1. Using (18) with the reference signal $x(t)$ as a cost function, a closed-form solution can be derived to solve the BD problem. For the details of the merits of the reference signal, see [6].

Researches using the idea of reference signals to solve the BSP problem, to our best knowledge, have been made by Adib et al. [1], [2], Rhioui [16], and Jelonnek et al. [6], [7]. Adib et al. [1], [2] have shown that the BSS for instantaneous mixtures can be achieved by maximizing $|D_{z_i x}|$ in (18) under the constraint $\sigma_{z_i}^2 = \sigma_{s_{\rho_i}}^2$ ($= 1$ if $\sigma_{s_i}^2 = 1$ for all $i = \overline{1, n}$), but they have not proposed any algorithm for achieving this idea, where $\sigma_{z_i}^2$ and $\sigma_{s_{\rho_i}}^2$ denote the variances of the output $z_i(t)$ and a source signal $s_{\rho_i}(t)$, respectively, and ρ_i is an integer $\{1, 2, \ldots, n\}$ such that the set $\{\rho_1, \rho_2, \ldots, \rho_n\}$ is a permutation of the set $\{1, 2, \ldots, n\}$. Rhioui et al. [16] have proposed quadratic MIMO contrast functions for the BSS with convolutive mixtures. In their method, the number of reference signals corresponds to the number of source signals that can be extracted. Moreover, they claimed that as a reference signal, it is a practical valid choice to choose a signal obtained by whitening the outputs of the MIMO convolved system. Jelonnek et al. [6], [7] have shown in the single-input case that, by the Lagrangian method, the maximization of $|D_{z_i x}|$ under $\sigma_{z_i}^2 = \sigma_{s_{\rho_i}}^2$ leads to a closed form expressed as a generalized eigenvector problem. In our case, $D_{z_i x}$ and $\sigma_{z_i}^2$ can be expressed in terms of the vector $\tilde{\boldsymbol{w}}_i$ as, respectively,

$$D_{z_i x} = \tilde{\boldsymbol{w}}_i^H \tilde{\tilde{\boldsymbol{B}}} \tilde{\boldsymbol{w}}_i, \tag{24}$$

$$\sigma_{z_i}^2 = \tilde{\boldsymbol{w}}_i^H \tilde{\tilde{\boldsymbol{R}}} \tilde{\boldsymbol{w}}_i, \tag{25}$$

where $\tilde{\boldsymbol{B}}$ is the $m \times m$ block matrix defined by

$$\tilde{\boldsymbol{B}} := \begin{bmatrix} \boldsymbol{B}_{11} & \boldsymbol{B}_{12} & \cdots & \boldsymbol{B}_{1m} \\ \boldsymbol{B}_{21} & \boldsymbol{B}_{22} & \cdots & \boldsymbol{B}_{2m} \\ \vdots & \vdots & \vdots & \vdots \\ \boldsymbol{B}_{m1} & \boldsymbol{B}_{m2} & \cdots & \boldsymbol{B}_{mm} \end{bmatrix} \tag{26}$$

whose (i, j)th block element $\boldsymbol{B}_{ij}$ is the matrix with the (l, r)th element $[\boldsymbol{B}_{ij}]_{lr}$ calculated by $\mathrm{cum}\{y_i^*(t-L_1-l+1), y_j(t-L_1-r+1), x^*(t), x(t)\}$ $(l, r = \overline{1, L})$ and $\tilde{\boldsymbol{R}} = E[\tilde{\boldsymbol{y}}^*(t)\tilde{\boldsymbol{y}}^T(t)]$ is the covariance matrix of m-block column vector $\tilde{\boldsymbol{y}}(t)$ defined by

$$\tilde{\boldsymbol{y}}(t) := \left[\boldsymbol{y}_1^T(t), \boldsymbol{y}_2^T(t), \ldots, \boldsymbol{y}_m^T(t)\right]^T \in \boldsymbol{C}^{mL}, \tag{27}$$

$$\boldsymbol{y}_j(t) := \left[y_j(t - L_1), y_j(t - L_1 - 1), \ldots, y_j(t - L_2)\right]^T \in \boldsymbol{C}^L, \qquad j = \overline{1, m}. \tag{28}$$

Therefore, by a similar method as used in [6], [7], the maximization of $|D_{z_i x}|$ under $\sigma_{z_i}^2 = \sigma_{s_{\rho_i}}^2$ leads to the following generalized eigenvector problem:

$$\tilde{\boldsymbol{B}}\tilde{\boldsymbol{w}}_i = \lambda\tilde{\boldsymbol{R}}\tilde{\boldsymbol{w}}_i. \tag{29}$$

Moreover, Jelonnek et al. have shown that the eigenvector corresponding to the maximum magnitude of the eigenvalues of $\tilde{\boldsymbol{R}}^\dagger\tilde{\boldsymbol{B}}$ becomes the solution of the blind equalization problem in [6], [7], which is referred to as an *eigenvector algorithm* (EVA). Note that since Jelonnek et al. have dealt with SISO-IIR channels or SIMO-IIR channels, the constructions of $\tilde{\boldsymbol{B}}$, $\tilde{\boldsymbol{w}}_i$, and $\tilde{\boldsymbol{R}}$ in (29) are different from those proposed in [6], [7]. In this paper, under the assumption that any reference filter $\boldsymbol{f}(z)$ is used, we want to show how the eigenvector algorithm (29) works for the BD of the MIMO-IIR channel (1).

To this end, we use the following equalities:

$$\tilde{\boldsymbol{R}} = \tilde{\boldsymbol{H}}^H\tilde{\boldsymbol{\Sigma}}\tilde{\boldsymbol{H}}, \tag{30}$$

$$\tilde{\boldsymbol{B}} = \tilde{\boldsymbol{H}}^H\tilde{\boldsymbol{\Lambda}}\tilde{\boldsymbol{H}}, \tag{31}$$

where $\tilde{\boldsymbol{\Sigma}}$ is the block diagonal matrix defined by

$$\tilde{\boldsymbol{\Sigma}} := \text{block-diag}\{\boldsymbol{\Sigma}_1, \boldsymbol{\Sigma}_2, \ldots, \boldsymbol{\Sigma}_n\}, \tag{32}$$

$$\boldsymbol{\Sigma}_i := \mathrm{diag}\{\ldots, \sigma_{s_i}^2, \sigma_{s_i}^2, \sigma_{s_i}^2, \ldots\}, \qquad i = \overline{1, n}, \tag{33}$$

and $\tilde{\boldsymbol{\Lambda}}$ is the block diagonal matrix defined by

$$\tilde{\boldsymbol{\Lambda}} := \text{block-diag}\{\boldsymbol{\Lambda}_1, \boldsymbol{\Lambda}_2, \ldots, \boldsymbol{\Lambda}_n\}, \tag{34}$$

$$\boldsymbol{\Lambda}_i := \mathrm{diag}\{\ldots, |a_i(-1)|^2\gamma_i, |a_i(0)|^2\gamma_i, |a_i(1)|^2\gamma_i, \ldots\}, \qquad i = \overline{1, n}. \tag{35}$$

As both $\tilde{\boldsymbol{\Sigma}}$ and $\tilde{\boldsymbol{\Lambda}}$ become diagonal, (30) and (31) show that the two matrices $\tilde{\boldsymbol{R}}$ and $\tilde{\boldsymbol{B}}$ are simultaneously diagonalizable. The proof of deriving (31) is given in Appendix A.

In the following theorem, we confine ourselves to the case $m = n$ for simplicity of discussion, although our results are expandable to the case $m > n$. Let the eigenvalues of the diagonal matrix $\tilde{\boldsymbol{\Sigma}}^{-1}\tilde{\boldsymbol{\Lambda}}$ be denoted by

$$\lambda_i(k) := |a_i(k)|^2\gamma_i/\sigma_{s_i}^2, \qquad i = \overline{1, n}, k \in Z. \tag{36}$$

Theorem 1. *Assume $L_1 = -\infty$ and $L_2 = \infty$, and suppose the following three conditions hold true:*

(T1-1) *There exist n integers $k_i \in Z$ such that the n eigenvalues $\lambda_i(k_i)$ are nonzero and distinct.*
(T1-2) *For each $i = \overline{1, n}$, $\lambda_i(k_i) \neq \lambda_i(k)$ for any $k \neq k_i$.*
(T1-3) *For each $i = \overline{1, n}$, $\lambda_i(k_i) \neq \lambda_j(k)$ for any $j \neq i$ and $k \in Z$.*

If the noise $\boldsymbol{n}(t)$ is absent in (1), then the n eigenvectors $\tilde{\boldsymbol{w}}_i$ corresponding to the n nonzero eigenvalues $\lambda_i(k_i)$ of $\tilde{\boldsymbol{R}}^{-1}\tilde{\boldsymbol{B}}$ become the vectors $\tilde{\boldsymbol{w}}_i$ satisfying (14).

Remark 2. The above three conditions can be replaced by a stronger condition that all the eigenvalues $\lambda_i(k)$ are distinct for $i = \overline{1, n}$ and $k \in Z$.

Proof. If $m = n$, the covariance matrix $\tilde{\boldsymbol{R}}$ is positive definite (see Appendix B or [11]). Therefore, based on (29), we consider the following eigenvector problem:

$$\tilde{\boldsymbol{R}}^{-1}\tilde{\boldsymbol{B}}\tilde{\boldsymbol{w}}_i = \lambda\tilde{\boldsymbol{w}}_i. \tag{37}$$

Then, from (30) and (31), (37) becomes

$$(\tilde{\boldsymbol{H}}^H\tilde{\boldsymbol{\Sigma}}\tilde{\boldsymbol{H}})^{-1}\tilde{\boldsymbol{H}}^H\tilde{\boldsymbol{\Lambda}}\tilde{\boldsymbol{H}}\tilde{\boldsymbol{w}}_i = \lambda\tilde{\boldsymbol{w}}_i. \tag{38}$$

It can be shown that $\tilde{\boldsymbol{H}}$ is nonsingular when $L_1 = -\infty$ and $L_2 = \infty$ (see Appendix C). Then (38) becomes

$$\tilde{\boldsymbol{H}}^{-1}\tilde{\boldsymbol{\Sigma}}^{-1}\tilde{\boldsymbol{\Lambda}}\tilde{\boldsymbol{H}}\tilde{\boldsymbol{w}}_i = \lambda\tilde{\boldsymbol{w}}_i. \tag{39}$$

Multiplying (39) by $\tilde{\boldsymbol{H}}$ from the left side, (39) becomes

$$\tilde{\boldsymbol{\Sigma}}^{-1}\tilde{\boldsymbol{\Lambda}}\tilde{\boldsymbol{H}}\tilde{\boldsymbol{w}}_i = \lambda\tilde{\boldsymbol{H}}\tilde{\boldsymbol{w}}_i. \tag{40}$$

Let

$$\tilde{\boldsymbol{g}}_i := \tilde{\boldsymbol{H}}\tilde{\boldsymbol{w}}_i, \tag{41}$$

then (40) becomes

$$\tilde{\boldsymbol{\Sigma}}^{-1}\tilde{\boldsymbol{\Lambda}}\tilde{\boldsymbol{g}}_i = \lambda\tilde{\boldsymbol{g}}_i. \tag{42}$$

Note that $\tilde{\boldsymbol{\Sigma}}^{-1}\tilde{\boldsymbol{\Lambda}}$ is a diagonal matrix with diagonal elements $\lambda_i(k)$, $i = \overline{1, n}$ and $k \in Z$, and that its diagonal elements are the eigenvalues of matrix $\tilde{\boldsymbol{R}}^{-1}\tilde{\boldsymbol{B}}$. From the three conditions **T1-1**, **T1-2** and **T1-3**, the n eigenvalues $\lambda_i(k_i)$ are nonzero and distinct, and for each $i = \overline{1, n}$, the eigenvalues $\lambda_i(k_i)$ are different from all the remainder eigenvalues of $\tilde{\boldsymbol{\Sigma}}^{-1}\tilde{\boldsymbol{\Lambda}}$. Therefore, the n nonzero eigenvectors $\tilde{\boldsymbol{g}}_i \neq 0$, $i = \overline{1, n}$, corresponding to the n nonzero eigenvalues $\lambda_i(k_i) \neq 0$, $i = \overline{1, n}$, obtained by (42), that is, the n nonzero eigenvectors $\tilde{\boldsymbol{w}}_i$, $i = \overline{1, n}$, corresponding to the n nonzero eigenvalues $\lambda_i(k_i) \neq 0$, $i = \overline{1, n}$, obtained by (37) become n solutions of the vectors $\tilde{\boldsymbol{w}}_i$ satisfying (14). □

Remark 3. In order to use Theorem 1, the reference signal $x(t)$ contains nonzero distinct contributions $a_i(k_i)$ from all source signals $s_i(t)$. This is a similar condition to that in [6], where $|a_i(k)|$ has only one maximum value with respect to $k \in Z$. Moreover, the parameters L_1 and L_2 of the deconvolver $\boldsymbol{W}(z)$ should be set to be $L_1 = -\infty$ and $L_2 = \infty$ in theory. However, it is impossible; hence, in practice we should approximate the infinite length to a finite length. In Section 4, we shall choose a length L for this approximation.

Remark 4. From Theorem 1, it can be seen that by all the n eigenvectors corresponding to the n nonzero eigenvalues $\lambda_i(k_i)$, all source signals can be separated from the output $\boldsymbol{y}(t)$. This is a novel result that has not been shown in the conventional researches. Moreover, it can be seen from (40) that even if the fourth-order cumulants γ_i have different signs for their values, the vector $\tilde{\boldsymbol{w}}_i$ satisfying (14) can be obtained. This fact will be confirmed by computer simulations in Section 4.

Remark 5. One can see that $\tilde{\boldsymbol{R}}^{-1}\tilde{\boldsymbol{B}}$ whose size is $nL \times nL$ has nL nonzero eigenvectors. Hence, it may be conjectured that an eigenvector except for the n eigenvectors $\tilde{\boldsymbol{w}}_i$ corresponding to the n nonzero eigenvalues $\lambda_i(k_i)$ may be used to recover a source signal.

3.2. Robust eigenvector algorithm

In the previous subsection, we assume that there is no noise in the output signals. In this subsection, we shall show an EVA such that the solutions (14) can be obtained, even if the noise $\boldsymbol{n}(t)$ satisfying the two assumptions, **A4** and **A5**, is presented in the output $\boldsymbol{y}(t)$. To this end, we introduce fourth-order cumulant matrices of the m-vector random process $\{\boldsymbol{y}(t)\}$ [19], which constitute a set of $m \times m$ block matrices $\boldsymbol{F}^{(4)}_{\boldsymbol{y},j,l}$, whose elements are defined by

$$\left[\boldsymbol{F}^{(4)}_{\boldsymbol{y},j,l}\right]_{[p,q]_{l_1 l_2}} = \mathrm{cum}\{y_q(t - L_1 - l_2 + 1),\, y_p^*(t - L_1 - l_1 + 1),\, y_j(t-l),\, y_j^*(t-l)\},$$
$$p, q, j = \overline{1, m},\, l_1, l_2 = \overline{1, L},\, l = \overline{L_1, L_2}, \tag{43}$$

where $[\cdot]_{[p,q]_{l_1 l_2}}$ denotes the (l_1, l_2)th element of the (p, q)th block matrix of the matrix $\boldsymbol{F}^{(4)}_{\boldsymbol{y},j,l}$. Then, we consider an $m \times m$ block matrix $\tilde{\boldsymbol{F}}$ expressed by

$$\tilde{\boldsymbol{F}} = \sum_{j=1}^{m} \sum_{l=L_1}^{L_2} \boldsymbol{F}^{(4)}_{\boldsymbol{y},j,l}. \tag{44}$$

It is shown by a simple calculation that (44) becomes

$$\tilde{\boldsymbol{F}} = \tilde{\boldsymbol{H}}^H \tilde{\boldsymbol{\Psi}} \tilde{\boldsymbol{H}}, \tag{45}$$

where $\tilde{\boldsymbol{\Psi}}$ is the diagonal matrix defined by

$$\tilde{\boldsymbol{\Psi}} := \text{block-diag}\{\boldsymbol{\Psi}_1, \boldsymbol{\Psi}_2, \ldots, \boldsymbol{\Psi}_n\}, \tag{46}$$

$$\boldsymbol{\Psi}_i := \text{diag}\{\ldots, \gamma_i \tilde{a}_i(-1), \gamma_i \tilde{a}_i(0), \gamma_i \tilde{a}_i(1), \ldots\}, \qquad i = \overline{1, n}, \tag{47}$$

$$\tilde{a}_i(k) := \sum_{j=1}^{m} \sum_{l=L_1}^{L_2} |h_{ji}(k-l)|^2, \qquad i = \overline{1, n}, \quad k \in Z. \tag{48}$$

We note from assumption **A1** that all $\tilde{a}_i(k)$ are positive, if $L_1 = -\infty$ and $L_2 = \infty$. As for (45), its derivation in the case of instantaneous mixtures is shown in the Appendix of [8]. Its extension to the case of convolutive mixtures is straightforward; hence, its details are omitted.

Here, as a constraint, we take the following value:

$$|D_{z_i y}| = \left| \sum_{j=1}^{m} \sum_{l=L_1}^{L_2} \text{cum}\{z_i(t), z_i^*(t), y_j(t-l), y_j^*(t-l)\} \right| = |\boldsymbol{w}_i^T \tilde{\boldsymbol{F}} \boldsymbol{w}_i|$$
$$= \left| \sum_{j=1}^{n} \gamma_j \sum_{k=-\infty}^{\infty} \tilde{a}_j(k)\, |g_{ri}(k)|^2 \right|. \tag{49}$$

Then, we consider solving the problem that the fourth-order cumulant $|D_{z_i x}|$ is maximized under the condition of $|D_{z_i y}| = |\gamma_{\rho_i} \tilde{a}_{\rho_i}(k)|$. Note that we may choose an appropriate positive value for $\tilde{a}_{\rho_i}(k)$, if its true value is not available. By the Lagrangian method, the following generalized eigenvector problem is derived from the above problem:

$$\tilde{\boldsymbol{B}} \tilde{\boldsymbol{w}}_i = \tilde{\lambda} \tilde{\boldsymbol{F}} \tilde{\boldsymbol{w}}_i. \tag{50}$$

From the following theorem, one can see that by solving the eigenvector problem of the matrix $\tilde{\boldsymbol{F}}^{-1}\tilde{\boldsymbol{B}}$, its solution provides the vectors $\tilde{\boldsymbol{w}}_i$ $(i = \overline{1, n})$ in (14). By the same reasoning as used in Theorem 1, we confine ourselves to the case $m = n$ in the following theorem. Let the eigenvalues of the diagonal matrix $\tilde{\boldsymbol{\Psi}}^{-1}\tilde{\boldsymbol{\Lambda}}$ be denoted by

$$\mu_i(k) := |a_i(k)|^2 / \tilde{a}_i(k), \qquad i = \overline{1, n}, k \in Z. \tag{51}$$

Theorem 2. *Assume $L_1 = -\infty$ and $L_2 = \infty$, and suppose the following three conditions hold true:*

(T2-1) *There exist n integers $k_i \in Z$ such that the n eigenvalues $\mu_i(k_i)$ are nonzero and distinct.*
(T2-2) *For each $i = \overline{1, n}$, $\mu_i(k_i) \neq \mu_i(k)$ for any $k \neq k_i$.*
(T2-3) *For each $i = \overline{1, n}$, $\mu_i(k_i) \neq \mu_j(k)$ for any $j \neq i$ and $k \in Z$.*

Then the n eigenvectors corresponding to n nonzero eigenvalues $\mu_i(k_i)$ of $\tilde{\boldsymbol{F}}^{-1}\tilde{\boldsymbol{B}}$ become the vectors $\tilde{\boldsymbol{w}}_i$, $i = \overline{1, n}$, satisfying (14).

Remark 6. The preceding three conditions can be replaced by a stronger condition that all the eigenvalues $\mu_i(k)$ are distinct for $i = \overline{1, n}$ and $k \in Z$.

Proof. It can be seen from Appendix B and (45) that if $m = n$, the matrix $\tilde{\boldsymbol{F}}$ is positive definite. Therefore, based on (50), we consider the following eigenvector problem:

$$\tilde{\boldsymbol{F}}^{-1}\tilde{\boldsymbol{B}}\tilde{\boldsymbol{w}}_i = \tilde{\lambda}\tilde{\boldsymbol{w}}_i. \tag{52}$$

By the similar method as used in (38) through (40), we obtain

$$\tilde{\boldsymbol{\Psi}}^{-1}\tilde{\boldsymbol{\Lambda}}\tilde{\boldsymbol{H}}\tilde{\boldsymbol{w}}_i = \tilde{\lambda}\tilde{\boldsymbol{H}}\tilde{\boldsymbol{w}}_i. \tag{53}$$

By using $\tilde{\boldsymbol{g}}_i$ in (41), (53) becomes

$$\tilde{\boldsymbol{\Psi}}^{-1}\tilde{\boldsymbol{\Lambda}}\tilde{\boldsymbol{g}}_i = \tilde{\lambda}\tilde{\boldsymbol{g}}_i. \tag{54}$$

We should note that $\tilde{\boldsymbol{\Psi}}^{-1}\tilde{\boldsymbol{\Lambda}}$ is a diagonal matrix with elements $\mu_i(k)$, $i = \overline{1,n}$ and $k \in Z$, and that its diagonal elements are the eigenvalues of $\tilde{\boldsymbol{F}}^{-1}\tilde{\boldsymbol{B}}$. From the three conditions **T2-1**, **T2-2**, and **T2-3**, therefore, the n nonzero eigenvectors $\tilde{\boldsymbol{g}}_i \neq 0$, $i = \overline{1,n}$, corresponding to the n nonzero eigenvalues $\mu_i(k_i) \neq 0$, $i = \overline{1,n}$, obtained by (54), that is, the n nonzero eigenvectors $\tilde{\boldsymbol{w}}_i$ $i = \overline{1,n}$, corresponding to the n nonzero eigenvalues $\mu_i(k_i) \neq 0$, $i = \overline{1,n}$, obtained by (52) become n solutions of the vectors $\tilde{\boldsymbol{w}}_i$ satisfying (14). □

Remark 7. The matrix $\tilde{\boldsymbol{F}}^{-1}\tilde{\boldsymbol{B}}$ consists of only fourth-order cumulants; thus, the eigenvectors derived from the matrix can be obtained with as little influence of Gaussian noise as possible (referred to as a REVA). This means that if one can estimate completely the same value as the theoretical one of $\tilde{\boldsymbol{F}}^{-1}\tilde{\boldsymbol{B}}$, one can obtain the same eigenvectors as if there were no noise in the output signal $\boldsymbol{y}(t)$. However, in fact, as there are estimation errors of the fourth-order cumulants, from the matrix $\tilde{\boldsymbol{F}}^{-1}\tilde{\boldsymbol{B}}$, we cannot obtain eigenvectors that are not affected by Gaussian noise (see Figure 3).

Remark 8. From the matrix $\tilde{\boldsymbol{\Psi}}^{-1}\tilde{\boldsymbol{\Lambda}}$ in (53), it can be seen that the fourth-order cumulants γ_i cancel each other in $\tilde{\boldsymbol{\Psi}}$ and $\tilde{\boldsymbol{\Lambda}}$. Therefore, the eigenvector algorithm (52) can be applied to the case where the signs of the fourth-order cumulants γ_i ($i = \overline{1,n}$) are different; that is, sub-Gaussian (the sign is minus $(-)$) and super-Gaussian (the sign is plus $(+)$) signals are treated as source signals.

Remark 9. The proposed EVAs in (37) and (52) are both closely related to the joint diagonalization of square matrices (e.g., [1], [3], [18]).

4. Computer simulations

To demonstrate the validity of the proposed method, many computer simulations were conducted. Some results are shown in this section. The unknown system $\boldsymbol{H}(z)$ was set to be an FIR channel with two inputs and two outputs, and we assumed that the length of the channel is three ($K = 3$), that is, the $\boldsymbol{H}^{(k)}$'s in (1) were set to be

$$\boldsymbol{H}(z) = \sum_{k=0}^{2} \boldsymbol{H}^{(k)} z^k = \begin{bmatrix} 1.00 + 0.15z + 0.10z^2 & 0.65 + 0.25z + 0.15z^2 \\ 0.50 - 0.10z + 0.20z^2 & 1.00 + 0.25z + 0.10z^2 \end{bmatrix}. \tag{55}$$

Gaussian noise $n_j(t)$ with variance $\sigma_{n_j}^2$ was included in the output $y_j(t)$ at various signal-to-noise ratio (SNR) levels. The SNR was considered at the output of the system $\boldsymbol{H}(z)$.

The matrix $\tilde{\boldsymbol{B}}$ in (37) and (52) was calculated using a moving average defined by

$$\begin{aligned} \tilde{\boldsymbol{B}}(t) &:= \beta_1 \tilde{\boldsymbol{B}}(t-1) + (1-\beta_1)\{|x(t)|^2 \boldsymbol{V}_1(t) \\ &-|x(t)|^2 \tilde{\boldsymbol{V}}_1(t) - x^*(t)\tilde{\boldsymbol{y}}^*(t)\tilde{\boldsymbol{v}}_1^T(t) - x(t)\tilde{\boldsymbol{y}}^*(t)\tilde{\boldsymbol{v}}_2^T(t)\}, \end{aligned} \tag{56}$$

where $\boldsymbol{V}_1(t) := \tilde{\boldsymbol{y}}^*(t)\tilde{\boldsymbol{y}}^T(t)$, $\tilde{\boldsymbol{V}}_1(t)$ is a moving average of $\boldsymbol{V}_1(t)$, calculated by

$$\tilde{\boldsymbol{V}}_1(t) = \beta_2 \tilde{\boldsymbol{V}}_1(t-1) + (1-\beta_2)\boldsymbol{V}_1(t), \tag{57}$$

and $\tilde{\boldsymbol{v}}_i$ $i = 1, 2$ are moving averages of $x(t)\tilde{\boldsymbol{y}}(t)$ and $x^*(t)\tilde{\boldsymbol{y}}(t)$, respectively, calculated by

$$\tilde{\boldsymbol{v}}_1(t) := \beta_2 \tilde{\boldsymbol{v}}_1(t-1) + (1-\beta_2)x(t)\tilde{\boldsymbol{y}}(t), \tag{58}$$

$$\tilde{\boldsymbol{v}}_2(t) := \beta_2 \tilde{\boldsymbol{v}}_2(t-1) + (1-\beta_2)x^*(t)\tilde{\boldsymbol{y}}(t). \tag{59}$$

Here $\tilde{\boldsymbol{y}}(t)$ is defined by (27) and (28), where m in (28) was 2 and the parameters L_1 and L_2 were set to be 0 and 9, respectively. Namely, the length L of the deconvolver $\boldsymbol{W}(z)$ was $L = 10$. The matrix $\tilde{\boldsymbol{F}}$ in (52) was calculated using a moving average defined by

$$\begin{aligned} \tilde{\boldsymbol{F}}(t) &:= \beta_1 \tilde{\boldsymbol{F}}(t-1) + (1-\beta_1)\{\tilde{\boldsymbol{V}\boldsymbol{V}}(t) - \tilde{\boldsymbol{V}}_1(t)\tilde{\boldsymbol{V}}_1^H(t) \\ &-\tilde{\boldsymbol{V}}_2(t)\tilde{\boldsymbol{V}}_2^H(t) - \mathrm{tr}\{\tilde{\boldsymbol{V}}_1(t)\}\tilde{\boldsymbol{V}}_1(t)\}, \end{aligned} \tag{60}$$

where $\tilde{\boldsymbol{V}\boldsymbol{V}}(t)$ is a moving average of $\boldsymbol{V}_1(t)\boldsymbol{V}_1^H(t)$ defined by

$$\tilde{\boldsymbol{V}\boldsymbol{V}}(t) = \beta_1 \tilde{\boldsymbol{V}\boldsymbol{V}}(t-1) + (1-\beta_1)\boldsymbol{V}_1(t)\boldsymbol{V}_1^H(t), \tag{61}$$

the matrix $\tilde{\boldsymbol{V}}_2(t)$ was calculated by

$$\tilde{\boldsymbol{V}}_2(t) = \beta_2 \tilde{\boldsymbol{V}}_2(t-1) + (1-\beta_2)\tilde{\boldsymbol{y}}(t)\tilde{\boldsymbol{y}}^T(t), \tag{62}$$

and $\mathrm{tr}\{X\}$ denotes the trace of the matrix X. As for the recurrence formula (60) of $\tilde{\boldsymbol{F}}(t)$, its derivation in the case of instantaneous mixtures is shown in Appendix B of [12]. Its extension to the case of convolutive mixtures is straightforward, and so its details are omitted. The matrix $\tilde{\boldsymbol{R}}$ was estimated by $\tilde{\boldsymbol{V}}_1(t)$ in (57), where in the following Examples 2 and 3, instead of β_2 in (57), the parameter β_1 in (56)

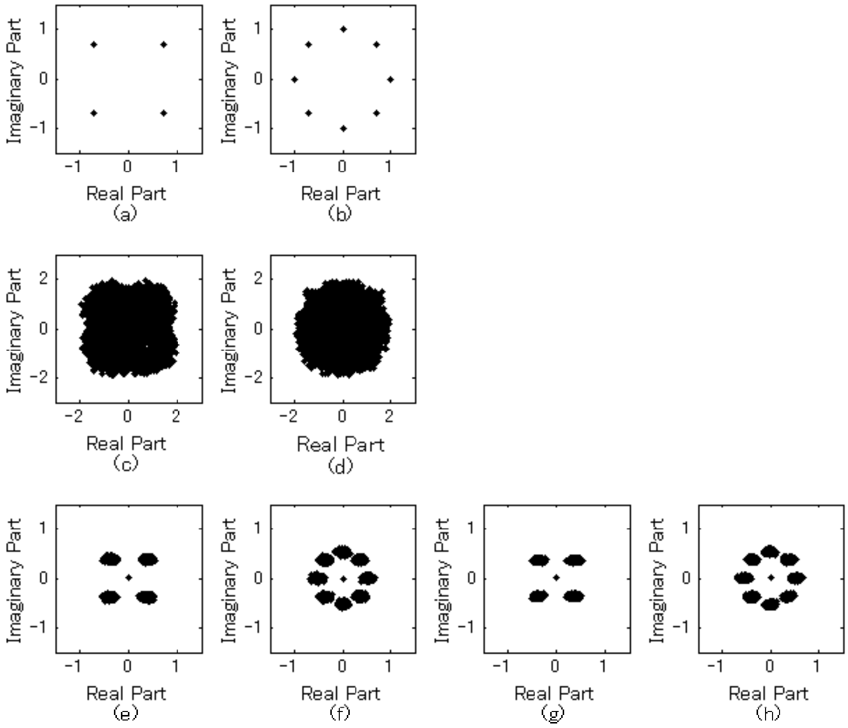

Figure 2. The source signals $s_i(t)$ ((a) and (b)), the channel outputs $y_j(t)$ ((c) and (d)), and the separated signals ((e) and (f), and (g) and (h)) obtained by the EVA and the REVA, respectively.

was used. Note that we have not yet obtained the selection rule for the parameters β_1 and β_2, but the values of the parameters are basically decided by the following rule: $1 > \beta_1 > \beta_2 > 0$. The L-column vectors $\boldsymbol{f}_j$, $j = 1, 2$ in (23) were set to be $\boldsymbol{f}_1 = [0, 1, 0, \ldots, 0]^T$, $\boldsymbol{f}_2 = [0, \ldots, 0]^T$, that is, $x(t) = y_1(t-1)$. Note that there are many ways of selecting for the reference filter $\boldsymbol{f}(z)$; the above filter provided the best performance from the simulations. As a measure of performance, we used the *multichannel intersymbol interference* ($\mathrm{M}_{\mathrm{ISI}}$) [5], [8]. The value of $\mathrm{M}_{\mathrm{ISI}}$ in the logarithmic (dB) scale becomes $-\infty$, if the $\tilde{\boldsymbol{g}}_i$'s in (14) are obtained, and hence a negative large value of M_{ISI} indicates proximity to the desired solution. When the M_{ISI} was calculated, the n eigenvectors of $\tilde{\boldsymbol{R}}^{-1}\tilde{\boldsymbol{B}}$ and $\tilde{\boldsymbol{F}}^{-1}\tilde{\boldsymbol{B}}$ were chosen such that the $\mathrm{M}_{\mathrm{ISI}}$ took as small values as possible.

Example 1. In this example, using a 4-quadrature amplitude modulation (QAM) and an 8-QAM as two source signals $s_1(t)$ and $s_2(t)$, we shall show that the proposed EVAs work well. The parameters β_1 in (56), (60), (61) and β_2 in (57)–(59), (62) were set to be $\beta_1 = 0.99999$ and $\beta_2 = 0.999$, respectively, with β_2 in (57) used by the EVA set to be 0.9995.

Figure 2 shows the source signals, the channel signals $y_j(t)$ $j = 1, 2$, and the separated signals obtained by the proposed EVA and REVA, where the SNR is 30 dB. The results were obtained by 10 iterations. That is, in each iteration, the estimations (56) through (62) were implemented by using 5,000 data samples, then, using the estimations, the eigenvectors of $\tilde{\boldsymbol{R}}^{-1}\tilde{\boldsymbol{B}}$ and $\tilde{\boldsymbol{F}}^{-1}\tilde{\boldsymbol{B}}$ were calculated. As the initial values of the estimations, the estimations obtained by previous iteration were used, and their initial values at the first iteration were given appropriately. From Figure 2, one can see that the proposed EVAs can achieve the BD successfully.

Example 2. In this example, the source signals $s_1(t)$ and $s_2(t)$ were a sub-Gaussian signal and a super-Gaussian signal, where the sub-Gaussian signal takes one of two values, -1 and 1 with equal probability 1/2 and the super-Gaussian signal takes -2, 2, and 0 with probabilities 1/8, 1/8, and 6/8, respectively. The parameters β_1 in (56), (60), (61) and β_2 in (57)–(59), (62) were set to be $\beta_1 = 0.999995$ and $\beta_2 = 0.9995$, respectively.

Figure 3 shows the results of performances for the proposed EVA and REVA when the SNR levels were respectively taken to be 10, 15, 20, 25, and 30 dB. The vertical and horizontal axes represent respectively M_{ISI} and SNR, where each M_{ISI} shown in Figure 3 was the average of the performance results obtained by 30 independent Monte Carlo runs. In each Monte Carlo run, the final eigenvectors of $\tilde{\boldsymbol{R}}^{-1}\tilde{\boldsymbol{B}}$ and $\tilde{\boldsymbol{F}}^{-1}\tilde{\boldsymbol{B}}$ were obtained by 10 iterative calculations, where in each iteration, $\tilde{\boldsymbol{F}}$, $\tilde{\boldsymbol{R}}$, and $\tilde{\boldsymbol{B}}$ were estimated by data samples in the following three cases: (Case 1) 5,000 samples, (Case 2) 10,000 samples, and (Case 3) 20,000 samples, and as their initial values, the estimated values obtained by the previous iteration were used, where the first iteration started with zeros. Note that a stopping rule of the iterative calculation for obtaining the final eigenvectors has not been obtained yet, but ten iterative calculations are enough to obtain the final eigenvectors.

It can be seen from Figure 3 that when the SNR level is more than about 20 dB, the EVA can provide better performances than the REVA; hence, the EVA may be more useful than the REVA at those SNR levels. On the other hand, the REVA is effective when the SNR level is less than about 20 dB, because as the number of data samples increases, the differences in performance between the REVA and the EVA become bigger. However, in the lower SNR case, that is, for less than 15 dB, the difference between the two performances of the REVA in Case 2 and Case 3 is small. Note that if the parameters β_1 in (61) and β_2 are changed, the performance of Case 3 becomes better than that of Case 2 (see the dashed lines in Figure 3), where we changed $\beta_2 = 0.9995$ to $\beta_2 = 0.999$ and the parameter β_1 in (61) was changed to β_2 ($= 0.999$). Therefore, in the lower SNR case (in Case 2, less than about 10 dB and in Case 3, less than about 15 dB), if the number of data samples used in the estimations is allowed to be large, we recommend that the parameter β_1 in (61) be changed to β_2 and that β_2 be changed appropriately. As for the sample size needed to estimate $\tilde{\boldsymbol{F}}$, $\tilde{\boldsymbol{R}}$, and $\tilde{\boldsymbol{B}}$, we consider that at least more than 5,000 data samples are needed.

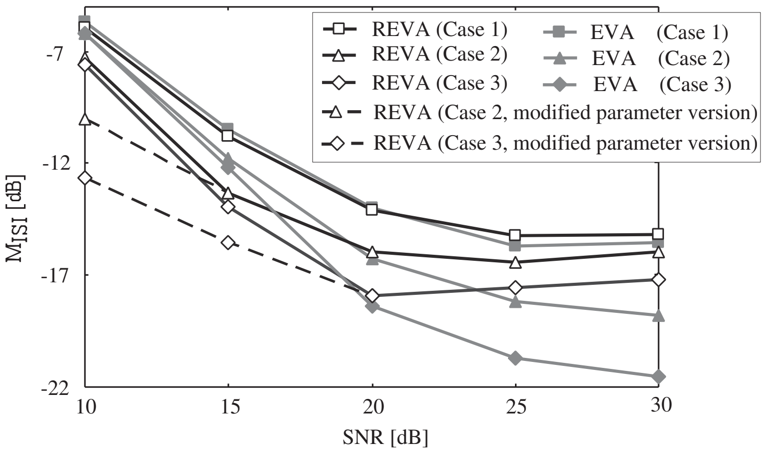

Figure 3. The performances of the REVA and the EVA with varying SNR levels for 5,000 samples (Case 1), 10,000 samples (Case 2), and 20,000 samples (Case 3).

Example 3. Finally, we shall show some results compared with the super-exponential method (SEM) (e.g., [5], [14], [17]) and the robust super-exponential method (RSEM) [8], [13], Those methods need deflation methods in order to separate source signals from their mixtures, and the RSEM has the same ability as the REVA. Our proposed EVAs can separate all source signals without the deflation method: This is one of their novel points. Therefore, we want to show some differences between the EVAs without the deflation method and the SEMs with it, in which the results are obtained by varying SNR levels.

As two methods compared with the EVA and the REVA, the SEM in [5] and the RSEM in [13] were used. As a deflation method, the one in [5] was used. The source signals used in the example were the sub-Gaussian and the super-Gaussian used in Example 2. The parameters of the estimations needed to implement the SEMs were chosen so that the SEMs gave the best performances.

Figure 4 shows the performance results. The horizontal and vertical axes represent the SNR levels (5 through 40 dB) and the values of M_{ISI}, where each M_{ISI} was the average of the results obtained by 30 independent Monte Carlo runs. In each Monte Carlo run, 10 iterations were implemented, and for each iteration, the estimations were implemented by using 5,000 data samples; then, using the estimations, the EVAs and SEMs were calculated iteratively. As shown in Figure 4, the proposed EVAs provide better performances than the SEMs. Note that, as the SNR level decreases, this tendency becomes remarkable. One of the reasons may be that the deflation method of the SEMs often fails to separate two source signals. However, we confirm that if the deflation method can be used successfully, the performances of the SEMs are similar to those of the EVAs.

Here, we show the computational complexities of the EVAs and SEMs using

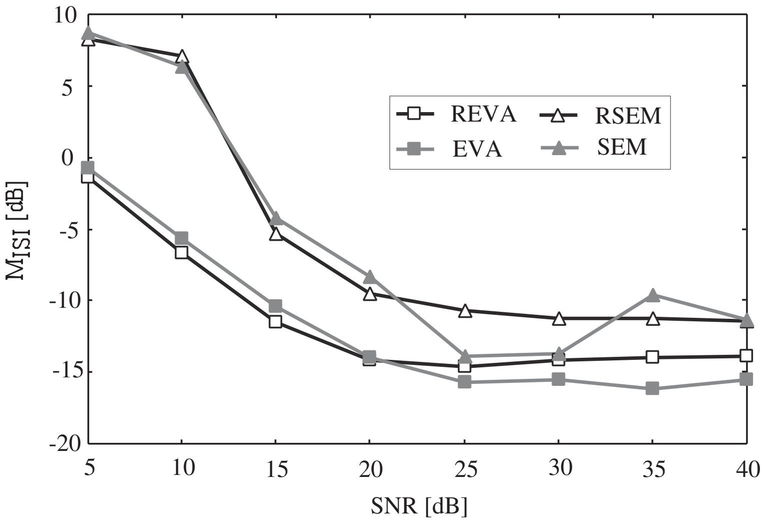

Figure 4. The comparison results of EVA, REVA, SEM, and RSEM with varying SNR levels.

Table 1. The flops and elapsed time of the EVAs and SEMs for one iteration for 5,000 samples

Method	Flops ($\log_{10}$FLO)	Time (sec.)
REVA	7.67	0.70
EVA	7.05	0.65
RSEM	7.89	0.91
SEM	6.95	0.82

the Flops (floating-point operation with MATLAB Ver. 5.2) defined in the logarithmic scale and the elapsed time for one iteration for 5,000 samples. Table 1 shows these results. The values of the elapsed time pertain to a Pentium 4, CPU 3.6GHz, 2GB RAM machine.

From all the results, we conclude that the proposed EVAs can work well for achieving the BD. In particular, the REVA shows better performances when the SNR level is less than about 20 dB, and when the SNR level is more than 20 dB, the EVA should be used to solve the BD problem. However, the SEMs can be used to solve the BD problem, if deflation methods work well. Finally, the computational complexity of the EVAs is almost the same as that of the SEMs.

5. Conclusions

We have proposed two types of EVA for solving the BD problem. By using reference signals, these EVAs are capable of separating source signals simultaneously from their mixtures. One of the EVAs is robust against Gaussian noise, which means that the EVA can be used to estimate the inverse of $\tilde{\boldsymbol{H}}$ with as little influence of Gaussian noise as possible. Computer simulations have demonstrated the effectiveness of the proposed EVAs.

The following problems have not been solved yet:

(i) How to select n eigenvectors corresponding to the eigenvalues (36) or (51) from the nL eigenvectors of $\tilde{\boldsymbol{R}}^{-1}\tilde{\boldsymbol{B}}$ or $\tilde{\boldsymbol{F}}^{-1}\tilde{\boldsymbol{B}}$.
(ii) The case of $m > n$.
(iii) The case of FIR channels.

As for (i), we do not have yet a complete scheme for selecting appropriate n eigenvectors from the nL eigenvectors of $\tilde{\boldsymbol{R}}^{-1}\tilde{\boldsymbol{B}}$ or $\tilde{\boldsymbol{F}}^{-1}\tilde{\boldsymbol{B}}$. However, we note that the condition stated in Remark 2 (or Remark 6) holds true except for pathological cases. Then, under the assumptions of $m = n$, $L_1 = -\infty$, and $L_2 = \infty$, $\tilde{\boldsymbol{H}}$ can be estimated from the matrix $\tilde{\boldsymbol{W}} = [\tilde{\boldsymbol{w}}_1, \ldots, \tilde{\boldsymbol{w}}_{nL}]$, that is, $\tilde{\boldsymbol{H}} \simeq \tilde{\boldsymbol{W}}^{-1}$ with $\tilde{\boldsymbol{W}}$ having the columns in correct order. If $\tilde{\boldsymbol{W}}$ has the columns in correct order, then it has the same structure as shown in (85) and (86). However, we have not yet found a correct scheme for attaining the objective. As for (ii), in the case of $m > n$, the matrices $\tilde{\boldsymbol{F}}^{\dagger}\tilde{\boldsymbol{B}}$ and $\tilde{\boldsymbol{R}}^{\dagger}\tilde{\boldsymbol{B}}$ are not full rank, that is, their ranks are nL $(< mL)$ [11]. We can conclude that the remaining $(m-n)L$ eigenvalues become zero. As for (iii), we have shown some results using computer simulations (see Section 4) in which the BD can be achieved by the proposed methods, but we have not yet obtained theoretical evidence to support the results. In the near future, we will find the solutions to these problems.

Appendix A. The proof for deriving (31)

From (27), we obtain

$$\tilde{\boldsymbol{y}}(t) = \tilde{\boldsymbol{H}}^T \tilde{\boldsymbol{s}}(t), \tag{63}$$

where $\tilde{\boldsymbol{s}}(t)$ is the column vector defined by

$$\tilde{\boldsymbol{s}}(t) := \left[\boldsymbol{s}_1^T(t), \boldsymbol{s}_2^T(t), \ldots, \boldsymbol{s}_n^T(t)\right]^T, \tag{64}$$

$$\boldsymbol{s}_j(t) := \left[\ldots, s_j(t+1), s_j(t), s_j(t-1), \ldots\right]^T, \qquad j = \overline{1, n}. \tag{65}$$

From (19), $x(t) = \tilde{\boldsymbol{a}}\tilde{\boldsymbol{s}}(t)$. Then, the block matrix $\tilde{\boldsymbol{B}}$ can be expressed as

$$\begin{aligned}\tilde{\boldsymbol{B}} = {} & E[\tilde{\boldsymbol{y}}^*(t)\tilde{\boldsymbol{y}}^T(t)x^*(t)x(t)] - E[\tilde{\boldsymbol{y}}^*(t)\tilde{\boldsymbol{y}}^T(t)]E[x^*(t)x(t)] \\ & -E[\tilde{\boldsymbol{y}}^*(t)x^*(t)]E[\tilde{\boldsymbol{y}}^T(t)x(t)] - E[\tilde{\boldsymbol{y}}^*(t)x(t)]E[\tilde{\boldsymbol{y}}^T(t)x^*(t)].\end{aligned} \tag{66}$$

From the assumption **A2**, we obtain

$$E[\tilde{\boldsymbol{y}}^*(t)\tilde{\boldsymbol{y}}^T(t)x^*(t)x(t)] = \tilde{\boldsymbol{H}}^H\boldsymbol{\Lambda}_1\tilde{\boldsymbol{H}},\ E[\tilde{\boldsymbol{y}}^*(t)\tilde{\boldsymbol{y}}^T(t)]E[x^*(t)x(t)] = \tilde{\boldsymbol{H}}^H\boldsymbol{\Lambda}_2\tilde{\boldsymbol{H}},$$
$$E[\tilde{\boldsymbol{y}}^*(t)x^*(t)]E[\tilde{\boldsymbol{y}}^T(t)x(t)] = \tilde{\boldsymbol{H}}^H\boldsymbol{\Lambda}_3\tilde{\boldsymbol{H}},\ E[\tilde{\boldsymbol{y}}^*(t)x(t)]E[\tilde{\boldsymbol{y}}^T(t)x^*(t)] = \tilde{\boldsymbol{H}}^H\boldsymbol{\Lambda}_2\tilde{\boldsymbol{H}}. \tag{67}$$

Here $\boldsymbol{\Lambda}_i$ $(i = 1, 2, 3)$ are diagonal matrices whose kth diagonal elements are $|a_i(k)|^2E[|s_i(t-k)|^4]$, $|a_i(k)|^2E[|s_i(t-k)|^2]^2$, and $|a_i(k)|^2E[s_i^*(t-k)^2]E[s_i(t-k)^2]$, respectively, $i = \overline{1, n}$, $k = \overline{-\infty, \infty}$. Therefore, from (66) and (67), we can find $\tilde{\boldsymbol{B}} = \tilde{\boldsymbol{H}}^H\tilde{\boldsymbol{\Lambda}}\tilde{\boldsymbol{H}}$, where $\tilde{\boldsymbol{\Lambda}} := \boldsymbol{\Lambda}_1 - 2\boldsymbol{\Lambda}_2 - \boldsymbol{\Lambda}_3$.

Appendix B. Proof of the positive definiteness of the covariance matrix $\tilde{\boldsymbol{R}}$

Let $\tilde{\boldsymbol{R}}$ be defined by $E[\tilde{\boldsymbol{y}}^*(t)\tilde{\boldsymbol{y}}^T(t)]$, using $\tilde{\boldsymbol{y}}(t) \in \boldsymbol{C}^{nL}$ in (27) and $\boldsymbol{y}_i(t) \in \boldsymbol{C}^L$, $i = \overline{1, n}$ in (28). Then it can be easily shown that $\tilde{\boldsymbol{R}}$ is positive semidefinite. Let

$$\bar{\boldsymbol{y}}(t) := \left[\dot{\boldsymbol{y}}^T(t-L_1), \dot{\boldsymbol{y}}^T(t-L_1-1), \ldots, \dot{\boldsymbol{y}}^T(t-L_2)\right]^T \in \boldsymbol{C}^{nL}, \tag{68}$$
$$\dot{\boldsymbol{y}}(t-l) := \left[y_1(t-l), y_2(t-l), \ldots, y_n(t-l)\right]^T \in \boldsymbol{C}^n, \qquad l = \overline{L_1, L_2}. \tag{69}$$

Then there exists a permutation matrix $\boldsymbol{P} \in \boldsymbol{C}^{nL\times nL}$ such that $\bar{\boldsymbol{y}}(t) = \boldsymbol{P}\tilde{\boldsymbol{y}}(t)$. Let $\bar{\boldsymbol{R}} := E[\bar{\boldsymbol{y}}^*(t)\bar{\boldsymbol{y}}^T(t)]$, then we have

$$\bar{\boldsymbol{R}} = \boldsymbol{P}\tilde{\boldsymbol{R}}\boldsymbol{P}^T. \tag{70}$$

It is clear that $\tilde{\boldsymbol{R}}$ is positive definite if and only if $\bar{\boldsymbol{R}}$ is positive definite.

Now suppose

$$\bar{\boldsymbol{\zeta}}^H\bar{\boldsymbol{R}}\bar{\boldsymbol{\zeta}} = 0, \tag{71}$$

where

$$\bar{\boldsymbol{\zeta}} := \left[\boldsymbol{\zeta}_{L_1}^T, \boldsymbol{\zeta}_{L_1+1}^T, \ldots, \boldsymbol{\zeta}_{L_2}^T\right]^T \in \boldsymbol{C}^{nL}, \tag{72}$$
$$\boldsymbol{\zeta}_l := [\zeta_{l1}, \zeta_{l2}, \ldots, \zeta_{ln}]^T \in \boldsymbol{C}^n, \qquad l = \overline{L_1, L_2}. \tag{73}$$

Since $\dot{\boldsymbol{y}}(t-L_1) = \sum_{k=-\infty}^{\infty}\boldsymbol{H}^{(k)}\boldsymbol{s}(t-L_1-k) = \sum_{k=-\infty}^{\infty}\boldsymbol{H}^{(k-L_1)}\boldsymbol{s}(t-k)$, we have

$$\begin{aligned}\bar{\boldsymbol{R}} &= E[\bar{\boldsymbol{y}}^*(t)\bar{\boldsymbol{y}}^T(t)]\\ &= \sum_{k=-\infty}^{\infty}\begin{bmatrix}\boldsymbol{H}^{(k-L_1)}\\ \vdots\\ \boldsymbol{H}^{(k-L_2)}\end{bmatrix}^*\boldsymbol{\Sigma}_\sigma\begin{bmatrix}\boldsymbol{H}^{(k-L_1)}\\ \vdots\\ \boldsymbol{H}^{(k-L_2)}\end{bmatrix}^T,\end{aligned} \tag{74}$$

where $\boldsymbol{\Sigma}_\sigma := \mathrm{diag}\{\sigma_{s_1}^2, \sigma_{s_2}^2, \ldots, \sigma_{s_n}^2\}$ whose diagonal elements are real values. Therefore, it follows from (71) and (74) that

$$\bar{\boldsymbol{\zeta}}^H \bar{\boldsymbol{R}} \bar{\boldsymbol{\zeta}} = \sum_{k=-\infty}^{\infty} \bar{\boldsymbol{\zeta}}^H \begin{bmatrix} \boldsymbol{H}^{(k-L_1)} \\ \vdots \\ \boldsymbol{H}^{(k-L_2)} \end{bmatrix}^* \boldsymbol{\Sigma}_\sigma \begin{bmatrix} \boldsymbol{H}^{(k-L_1)} \\ \vdots \\ \boldsymbol{H}^{(k-L_2)} \end{bmatrix}^T \bar{\boldsymbol{\zeta}} = 0. \tag{75}$$

Hence from (75), we have

$$\bar{\boldsymbol{\zeta}}^T \begin{bmatrix} \boldsymbol{H}^{(k-L_1)} \\ \vdots \\ \boldsymbol{H}^{(k-L_2)} \end{bmatrix} \boldsymbol{\Sigma}_\sigma^{\frac{1}{2}} = \boldsymbol{0}, \qquad k \in Z, \tag{76}$$

which means

$$\bar{\boldsymbol{\zeta}}^T \begin{bmatrix} \boldsymbol{H}^{(k-L_1)} \\ \vdots \\ \boldsymbol{H}^{(k-L_2)} \end{bmatrix} = \boldsymbol{0}, \qquad k \in Z, \tag{77}$$

because $\boldsymbol{\Sigma}_\sigma^{\frac{1}{2}} > 0$, where $\boldsymbol{\Sigma}_\sigma^{\frac{1}{2}} = \mathrm{diag}\{\sigma_{s_1}, \sigma_{s_2}, \ldots, \sigma_{s_n}\}$. It follows from (72) and (76) that

$$[\boldsymbol{\zeta}_{L_1}^T, \boldsymbol{\zeta}_{L_1+1}^T, \ldots \boldsymbol{\zeta}_{L_2}^T] \begin{bmatrix} \boldsymbol{H}^{(k-L_1)} \\ \boldsymbol{H}^{(k-L_1-1)} \\ \vdots \\ \boldsymbol{H}^{(k-L_2)} \end{bmatrix} z^k = \boldsymbol{0} \quad \text{for } z \in C, k \in Z, \tag{78}$$

which implies

$$\sum_{k=-\infty}^{\infty} \sum_{l=L_1}^{L_2} \boldsymbol{\zeta}_l^T z^l \boldsymbol{H}^{(k)} z^k = \boldsymbol{0} \quad \text{for } z \in C, \tag{79}$$

which is equivalent to

$$\boldsymbol{\zeta}^T(z) \boldsymbol{H}(z) = \boldsymbol{0} \quad \text{for } z \in C, \tag{80}$$

where $\boldsymbol{\zeta}(z) := \boldsymbol{\zeta}_{L_1} z^{L_1} + \boldsymbol{\zeta}_{L_1+1} z^{L_1+1} + \cdots + \boldsymbol{\zeta}_{L_2} z^{L_2} \in \boldsymbol{C}^n$. Because $\boldsymbol{H}(z)$ has full rank on the unit circle $|z| = 1$, from (80), we have $\boldsymbol{\zeta}(z) = \boldsymbol{0}$ on the unit circle. This is equivalent to $\boldsymbol{\zeta}(z) = \boldsymbol{0}$ for any $z \in C$. Therefore, we have $\boldsymbol{\zeta}_i = \boldsymbol{0}$ for $i = \overline{L_1, L_2}$. Hence, $\bar{\boldsymbol{\zeta}} = \boldsymbol{0}$, and then we get $\mathrm{Ker}\,\bar{\boldsymbol{R}} = \{0\}$, where the symbol $\mathrm{Ker}\,\bar{\boldsymbol{R}}$ denotes the kernel of matrix $\bar{\boldsymbol{R}}$.

Appendix C. Existence of the inverse of $\tilde{H}$, where $L_1 = -\infty$, $L_2 = \infty$

Let

$$\tilde{\boldsymbol{y}} := \tilde{\boldsymbol{H}}^T \tilde{\boldsymbol{s}}, \tag{81}$$

where

$$\begin{aligned}
\tilde{\boldsymbol{y}} &:= [\boldsymbol{y}_1^T, \boldsymbol{y}_2^T, \ldots, \boldsymbol{y}_n^T]^T, \\
\boldsymbol{y}_i &:= [\ldots, y_i(-1), y_i(0), y_i(1), \ldots]^T, \qquad i = \overline{1, n}, \\
\tilde{\boldsymbol{s}} &:= [\boldsymbol{s}_1^T, \boldsymbol{s}_2^T, \ldots, \boldsymbol{s}_n^T]^T, \\
\boldsymbol{s}_i &:= [\ldots, s_i(-1), s_i(0), s_i(1), \ldots]^T, \qquad i = \overline{1, n}.
\end{aligned}$$

In the time domain, (81) is equivalent to (1). Furthermore, applying Wiener's inversion theorem [4] to $\boldsymbol{H}(z)$, the assumption **A1** ensures the existence of its stable inverse $\boldsymbol{H}^{-1}(z)$.

Here, we may write (1) as

$$\boldsymbol{y}(t) = \boldsymbol{H}(z)\boldsymbol{s}(t), \tag{82}$$

and then putting $\boldsymbol{W}(z) = \boldsymbol{H}^{-1}(z)$, we have from (82)

$$\boldsymbol{W}(z)\boldsymbol{y}(t) = \sum_{k=-\infty}^{\infty} \boldsymbol{W}^{(k)}\boldsymbol{y}(t-k) = \boldsymbol{s}(t), \tag{83}$$

where the second equality comes from

$$\sum_{\tau=-\infty}^{\infty} \boldsymbol{W}^{(\tau)}\boldsymbol{H}^{(k-\tau)} = \delta(k)\boldsymbol{I}. \tag{84}$$

Let

$$\tilde{\boldsymbol{W}} := \begin{bmatrix} \boldsymbol{W}_{11} & \boldsymbol{W}_{12} & \cdots & \boldsymbol{W}_{1n} \\ \boldsymbol{W}_{21} & \boldsymbol{W}_{22} & \cdots & \boldsymbol{W}_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ \boldsymbol{W}_{n1} & \boldsymbol{W}_{n2} & \cdots & \boldsymbol{W}_{nn} \end{bmatrix}, \tag{85}$$

$$\left[\boldsymbol{W}_{ij}\right]_{lr} := w_{ji}(l-r), \qquad l \in Z, r \in Z. \tag{86}$$

Then (83) is equivalent to

$$\tilde{\boldsymbol{W}}^T \tilde{\boldsymbol{y}} = \tilde{\boldsymbol{s}}. \tag{87}$$

(81) and (87) mean $\tilde{\boldsymbol{W}}^T\tilde{\boldsymbol{H}}^T = \boldsymbol{I}$ and $\tilde{\boldsymbol{H}}^T\tilde{\boldsymbol{W}}^T = \boldsymbol{I}$. Therefore, the inverse of $\tilde{\boldsymbol{H}}^T$ exists.

Acknowledgments

The authors thank Shuhei Kobayashi, Suguru Noguchi, and Takehide Maeda for their cooperation in the computer simulations in this paper.

References

[1] A. Adib, E. Moreau, and B. Aboutajdine, Blind source separation by simultaneous generalized referenced contrasts diagonalization, in *Proc. of Int. Sympo. on ICA and BSS*, pp. 657–661, 2003.

[2] A. Adib, E. Moreau, and D. Aboutajdine, Source separation contrasts using a reference signal, *IEEE Signal Processing Letters*, vol. 11, no. 3, pp. 312–315, 2004.

[3] J.-F. Cardoso and A. Souloumiac, Jacobi angles for simultaneous diagonalization, *SIAM J. Matrix Anal. Appl.*, vol. 17, pp. 161–164, 1996.

[4] Y. Inouye and K. Hirano, Cumulant-based blind identification of linear multi-input-multi-output systems driven by colored inputs, *IEEE Trans. Signal Processing*, vol. 45, no. 6, pp. 1543–1552, 1997.

[5] Y. Inouye and K. Tanebe, Super-exponential algorithms for multichannel blind deconvolution, *IEEE Trans. Signal Processing*, vol. 48, no. 3, pp. 881–888, 2000.

[6] B. Jelonnek and K. D. Kammeyer, A closed-form solution to blind equalization, *Signal Processing*, 36(3), pp. 251–259, 1994.

[7] B. Jelonnek, D. Boss, and K. D. Kammeyer, Generalized eigenvector algorithm for blind equalization, *Signal Processing*, 61(3), pp. 237–264, 1997.

[8] M. Kawamoto, K. Kohno, and Y. Inouye, Robust super-exponential methods for deflationary blind source separation of instantaneous mixtures, *IEEE Trans. Signal Processing*, vol. 53, no. 5, pp. 1933–1937, 2005.

[9] M. Kawamoto, K. Kohno, and Y. Inouye, Eigenvector algorithms using reference signals, in *Proc. of Int. Conf. on Acoustics, Speech, and Signal Process.*, vol. V, pp. 841–844, 2006.

[10] M. Kawamoto, K. Kohno, and Y. Inouye, Eigenvector algorithms using reference signals for blind source separation of instantaneous mixtures, in *Proc. of Int. Sympo. on Circuits and Systems*, pp. 4191–4194, 2006.

[11] K. Kohno, Y. Inouye, M. Kawamoto, and T. Okamoto, Adaptive super-exponential algorithms for blind deconvolution of MIMO systems, in *Proc. of Int. Sympo. on Circuits and Systems*, vol. V, pp. 680–683, 2004.

[12] K. Kohno, M. Kawamoto, A. K. Nandi, and Y. Inouye, Robust blind equalization algorithms based on the constrained maximization of a fourth-order cumulant function, *IEICE Transactions on Fundamentals*, vol. E89-A, no. 5, pp. 1495–1499, 2006.

[13] K. Kohno, Y. Inouye, and M. Kawamoto, Robust super-exponential methods for blind equalization of MIMO-IIR systems, in *Proc. of Int. Conf. on Acoustics, Speech, and Signal Process.*, vol. V, pp. 661–664, 2006.

[14] M. Martone, An adaptive algorithm for antenna array low-rank processing in cellular TDMA base stations, *IEEE Trans. Commum.*, vol. 46, no. 5, pp. 627–643, 1998.

[15] L. Parra and P. Sajda, Blind source separation via generalized eigenvalue decomposition, *Journal of Machine Learning*, no. 4, pp, 1261–1269, 2003.

[16] S. Rhioui, M. Castella, and E. Moreau, Quadratic MIMO contrast functions for blind source separation in a convolutive contest, in *Proc. of Int. Sympo. on ICA and BSS*, LNCS 3889, Springer-Verlag, Berlin Heidelberg, 2006, pp. 230–237.

[17] O. Shalvi and E. Weinstein, Super-exponential methods for blind deconvolution, *IEEE Trans. Information Theory*, vol. 39, no. 2, pp. 504–519, 1993.

[18] F. Theis and Y. Inouye, On the use of joint diagonalization in blind signal processing, in *Proc. of Int. Sympo. on Circuits and Systems*, pp. 3586–3589, 2006

[19] L. Tong, Y. Inouye, and R.-W. Liu, Waveform-preserving blind estimation of multiple independent sources, *IEEE Trans. Signal Processing*, vol. 41, no. 7, pp. 2461–2470, 1993.