

Blind Deconvolution of MIMO-IIR Systems: A Two-Stage EVA*

Mitsuru Kawamoto¹, Yujiro Inouye², and Kiyotaka Kohno³

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

² Department of Electronic and Control Systems Engineering, Shimane University,
1060 Nishikawatsu, Matsue, 690-8504, Japan

³ Department of Electronic Control Engineering, Yonago National College of
Technology, 4448 Hikona, Yonago, Tottori 683-8502, Japan
m.kawamoto@aist.go.jp, inouye@riko.shimane-u.ac.jp, kohno@yonago-k.ac.jp

Abstract. This paper deals with a blind deconvolution (DB) problem for multiple-input multiple-output infinite impulse response (MIMO-IIR) systems. To solve this problem, we propose an eigenvector algorithm (EVA). In the proposed EVA, two kinds of EVAs are merged so as to give a good performance: One is an EVA and the other is a Robust EVA (REVA) which works with as little sensitive to Gaussian noise as possible. Owing to this combination, two drawbacks of the conventional EVAs can be overcome. Simulation results show the validity of the proposed EVA.

Keywords: Independent component analysis, Blind deconvolution, Eigenvector algorithms, MIMO-IIR, Reference systems.

1 Introduction

This paper deals with a blind deconvolution (BD) problem for a multiple-input and multiple-output (MIMO) infinite impulse response (IIR) systems. To solve this problem, we use eigenvector algorithms (EVAs) [5,6,13]. The first proposal of the EVA was done by Jelonnek et al. [5]. They have proposed the EVA for solving blind equalization (BE) problems of single-input single-output (SISO) systems or single-input multiple-output (SIMO) systems. In [13], several procedures for the blind source separation (BSS) of instantaneous mixtures, using the generalized eigenvalue decomposition (GEVD), have been introduced. Recently, the authors have proposed the EVAs which can solve blind source separation (BSS) problems in the case of MIMO static systems (instantaneous mixtures) [7,8]. Moreover, based on the idea in [7], an EVA was derived for MIMO-IIR systems (convolutive mixtures) [9].

In the EVAs in [7,8,9], an idea of using reference signals was adopted. Researches applying this idea to solving blind signal processing (BSP) problems,

* Parts of the results in this paper were presented at IEEE Int. Conf. on Acoustics, Speech and Signal Processing, April 2007.

such as the BD, the BE, the BSS, and so on, have been made by Jelonnek et al. (e.g., [5]), Adib et al. (e.g., [1]), Rhioui et al. [14], and Castella, et al. [2]. Jelonnek et al. have shown in the single-input case that by the Lagrangian method, the maximization of a contrast function leads to a closed-form solution expressed as a generalized eigenvector problem, which is referred to as an eigenvector algorithm (EVA). Adib et al. have shown that the BSS for instantaneous mixtures can be achieved by maximizing a contrast function, but they have not proposed any algorithm for achieving this idea. Rhioui et al. [14] and Castella et al. [2] have proposed quadratic MIMO contrast functions for the BSS with convolutive mixtures, and have proposed an algorithm for extracting one source signal using a "fixed point"-like method. However, they have not presented a theoretical proof for the convergence of their proposed algorithm. In order to recover all source signals, in [14], the reference signals corresponding to the number of source signals which can be extracted were used, and in [2], a deflation approach was used, in which for each deflation, a different reference signal was used. The EVA in [9] can work so as to recover simultaneously all source signals using only one reference signal. However, the EVA has different performances for a different choice of the reference signal. Moreover, the conventional EVAs, e.g., [2,9], are sensitive to Gaussian noise.

In this paper, based on [8], we extend the EVA in [9] so as to work with as little influence of Gaussian noise as possible, which is referred to as a robust EVA (REVA). Then reference signals are chosen by utilizing the idea in [2], in order to overcome the drawback of the EVA in [9]. However this choice causes to use deflation methods. Since we want to show an algorithm without using the deflation method, then the following procedure is proposed as a two-stage EVA: 1) the EVA in [9] is executed with several iterations so that the solutions achieving the BD can be roughly found. As the next stage, 2) the REVA is executed using reference signals obtained by the scheme in [2], which are defined as appropriately chosen outputs of the deconvolvers with the filters obtained in Stage 1). Since the filters obtained in Stage 1) are linearly independent, the REVA with the reference signals based on the filters does not need deflation methods. Using this two-stage EVA, we try to overcome two drawbacks of the conventional EVAs, that is, the performances of the EVA are

- i) sensitive to Gaussian noise.
- ii) sensitive to choosing reference signals.

Simulation results show the validity of the proposed algorithm.

The present paper uses the following notation: Let Z denote the set of all integers. Let C denote the set of all complex numbers. Let C^n denote the set of all n -column vectors with complex components. Let $C^{m \times n}$ denote the set of all $m \times n$ matrices with complex components. The superscripts T , $*$, and H denote, respectively, the transpose, the complex conjugate, and the complex conjugate transpose (Hermitian) of a matrix. The symbols $\text{block-diag}\{\dots\}$ and $\text{diag}\{\dots\}$ denote respectively a block diagonal and a diagonal matrices with the

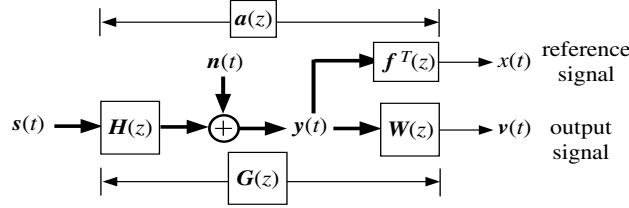


Fig. 1. The composite system of an unknown system and a deconvolver, and a reference system

block diagonal and the diagonal elements $\{\cdot\cdot\cdot\}$. The symbol $\text{cum}\{x_1, x_2, x_3, x_4\}$ denotes the fourth-order cumulant of x_i 's. Let $i = \overline{1, n}$ stand for $i = 1, 2, \dots, n$.

2 Problem Formulation and Assumptions

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

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

where $\mathbf{s}(t)$ is an n -column vector of input (or source) signals, $\mathbf{y}(t)$ is an m -column vector of system outputs, $\mathbf{n}(t)$ is an m -column vector of Gaussian noises, and $\{\mathbf{H}^{(k)}\}$ is an $m \times n$ impulse response matrix sequence. The transfer function of the system is defined by $\mathbf{H}(z) = \sum_{k=-\infty}^{\infty} \mathbf{H}^{(k)} z^k, z \in C$.

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

$$\mathbf{v}(t) = \sum_{k=-\infty}^{\infty} \mathbf{W}^{(k)} \mathbf{y}(t-k) = \sum_{k=-\infty}^{\infty} \mathbf{G}^{(k)} \mathbf{s}(t-k) + \sum_{k=-\infty}^{\infty} \mathbf{W}^{(k)} \mathbf{n}(t-k), \quad (2)$$

where $\{\mathbf{G}^{(k)}\}$ is the impulse response matrix sequence of $\mathbf{G}(z) := \mathbf{W}(z)\mathbf{H}(z)$, which is defined by $\mathbf{G}(z) = \sum_{k=-\infty}^{\infty} \mathbf{G}^{(k)} z^k, z \in C$. The cascade connection of the unknown system and the deconvolver is illustrated in Fig. 1.

Here, we put the following assumptions on the system, the source signals, the deconvolver, and the noises.

A1) The transfer function $\mathbf{H}(z)$ 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 $\{\mathbf{s}(t)\}$ 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

$$\gamma_i = \text{cum}\{s_i(t), s_i(t), s_i^*(t), s_i^*(t)\} \neq 0. \quad (3)$$

A3) The deconvolver $\mathbf{W}(z)$ is an FIR system of sufficient length L so that the truncation effect can be ignored.

A4) The noise sequence $\{\mathbf{n}(t)\}$ 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 $\{\mathbf{n}(t)\}$ and $\{\mathbf{s}(t)\}$ are mutually statistically independent.

Under **A3)**, the impulse response $\{\mathbf{G}^{(k)}\}$ of the cascade system is given by

$$\mathbf{G}^{(k)} := \sum_{\tau=L_1}^{L_2} \mathbf{W}^{(\tau)} \mathbf{H}^{(k-\tau)}, \quad k \in Z, \quad (4)$$

where the length $L := L_2 - L_1 + 1$ is taken to be sufficiently large. In a vector form, (4) can be written as

$$\tilde{\mathbf{g}}_i = \tilde{\mathbf{H}} \tilde{\mathbf{w}}_i, \quad i = \overline{1, n}, \quad (5)$$

where $\tilde{\mathbf{g}}_i$ is the column vector consisting of the i th output impulse response of the cascade system defined by $\tilde{\mathbf{g}}_i := [\mathbf{g}_{i1}^T, \mathbf{g}_{i2}^T, \dots, \mathbf{g}_{in}^T]^T$,

$$\mathbf{g}_{ij} := [\dots, g_{ij}(-1), g_{ij}(0), g_{ij}(1), \dots]^T, \quad j = \overline{1, n} \quad (6)$$

where $g_{ij}(k)$ is the (i, j) th element of matrix $\mathbf{G}^{(k)}$, and $\tilde{\mathbf{w}}_i$ is the mL -column vector consisting of the tap coefficients (corresponding to the i th output) of the deconvolver defined by $\tilde{\mathbf{w}}_i := [\mathbf{w}_{i1}^T, \mathbf{w}_{i2}^T, \dots, \mathbf{w}_{im}^T]^T \in \mathbf{C}^{mL}$,

$$\mathbf{w}_{ij} := [w_{ij}(L_1), w_{ij}(L_1 + 1), \dots, w_{ij}(L_2)]^T \in \mathbf{C}^L, \quad (7)$$

$j = \overline{1, m}$, where $w_{ij}(k)$ is the (i, j) th element of matrix $\mathbf{W}^{(k)}$, and $\tilde{\mathbf{H}}$ is the $n \times m$ block matrix whose (i, j) th block element \mathbf{H}_{ij} is the matrix (of L columns and possibly infinite number of rows) with the (l, r) th element $[\mathbf{H}_{ij}]_{lr}$ defined by $[\mathbf{H}_{ij}]_{lr} := h_{ji}(l - r)$, $l = 0, \pm 1, \pm 2, \dots$, $r = \overline{L_1, L_2}$, where $h_{ij}(k)$ is the (i, j) th element of the matrix $\mathbf{H}^{(k)}$.

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

$$[\tilde{\mathbf{g}}_1, \dots, \tilde{\mathbf{g}}_n] = \tilde{\mathbf{H}}[\tilde{\mathbf{w}}_1, \dots, \tilde{\mathbf{w}}_n] = [\tilde{\boldsymbol{\delta}}_1, \dots, \tilde{\boldsymbol{\delta}}_n] \mathbf{P}, \quad (8)$$

where \mathbf{P} is an $n \times n$ permutation matrix, and $\tilde{\boldsymbol{\delta}}_i$ is the n -block column vector defined by $\tilde{\boldsymbol{\delta}}_i := [\boldsymbol{\delta}_{i1}^T, \boldsymbol{\delta}_{i2}^T, \dots, \boldsymbol{\delta}_{in}^T]^T$, $i = \overline{1, n}$, $\boldsymbol{\delta}_{ij} := \hat{\boldsymbol{\delta}}_i$, for $i = j$, otherwise $(\dots, 0, 0, 0, \dots)^T$. Here, $\hat{\boldsymbol{\delta}}_i$ is the column vector (of infinite elements) whose r th element $\hat{\delta}_i(r)$ is given by $\hat{\delta}_i(r) = d_i \delta(r - k_i)$, 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 EVAs with Reference Signals

Jelonnek et al. [5] have shown in the single-input case that from the following problem, that is,

$$\text{Maximize } D_{v_i x} = \text{cum}\{v_i(t), v_i^*(t), x(t), x^*(t)\} \quad \text{under } \sigma_{v_i}^2 = \sigma_{s_{\rho_i}}^2, \quad (9)$$

a closed-form solution expressed as a generalized eigenvector problem can be led by the Lagrangian method, where $\sigma_{v_i}^2$ and $\sigma_{s_{\rho_i}}^2$ denote the variances of the output $v_i(t)$ and a source signal $s_{\rho_i}(t)$, respectively, and ρ_i is one of integers $\{1, 2, \dots, n\}$ such that the set $\{\rho_1, \rho_2, \dots, \rho_n\}$ is a permutation of the set $\{1, 2, \dots, n\}$, $v_i(t)$ is the i th element of $\mathbf{v}(t)$ in (2), the reference signal $x(t)$ is given by $\mathbf{f}^T(z)\mathbf{y}(t)$ using an appropriate filter $\mathbf{f}(z)$ (see Fig. 1). The filter $\mathbf{f}(z)$ is called a *reference system*. Let $\mathbf{a}(z) := \mathbf{H}^T(z)\mathbf{f}(z) = [a_1(z), a_2(z), \dots, a_n(z)]^T$, then $x(t) = \mathbf{f}^T(z)\mathbf{H}(z)\mathbf{s}(t) = \mathbf{a}^T(z)\mathbf{s}(t)$. The element $a_i(z)$ of the filter $\mathbf{a}(z)$ is defined as $a_i(z) = \sum_{k=-\infty}^{\infty} a_i(k)z^k$ and the reference system $\mathbf{f}(z)$ is an m -column vector whose elements are $f_j(z) = \sum_{k=L_1}^{L_2} f_j(k)z^k$, $j = \overline{1, m}$.

In our case, $D_{v_i x}$ and $\sigma_{v_i}^2$ can be expressed in terms of the vector $\tilde{\mathbf{w}}_i$ as, respectively, $D_{v_i x} = \tilde{\mathbf{w}}_i^H \tilde{\mathbf{B}} \tilde{\mathbf{w}}_i$ and $\sigma_{v_i}^2 = \tilde{\mathbf{w}}_i^H \tilde{\mathbf{R}} \tilde{\mathbf{w}}_i$, where $\tilde{\mathbf{B}}$ is the $m \times m$ block matrix whose (i, j) th block element \mathbf{B}_{ij} is the matrix with the (l, r) th element $[\mathbf{B}_{ij}]_{lr}$ calculated by $\text{cum}\{y_i^*(t-L_1-l+1), y_j(t-L_1-r+1), x^*(t), x(t)\}$ ($l, r = \overline{1, L}$) and $\tilde{\mathbf{R}} = E[\tilde{\mathbf{y}}^*(t)\tilde{\mathbf{y}}^T(t)]$ is the covariance matrix of m -block column vector $\tilde{\mathbf{y}}(t)$ defined by

$$\tilde{\mathbf{y}}(t) := [\mathbf{y}_1^T(t), \mathbf{y}_2^T(t), \dots, \mathbf{y}_m^T(t)]^T \in \mathbf{C}^{mL}, \tag{10}$$

where $\mathbf{y}_j(t) := [y_j(t-L_1), y_j(t-L_1-1), \dots, y_j(t-L_2)]^T \in \mathbf{C}^L$, $j = \overline{1, m}$. It follows from (10) that $\tilde{\mathbf{y}}(t)$ is expressed as $\tilde{\mathbf{y}}(t) = \mathbf{D}_c(z)\mathbf{y}(t)$, where $\mathbf{D}_c(z)$ is an $mL \times m$ converter (consisting of m identical delay chains each of which has L delay elements when $L_1 = 1$) defined by $\mathbf{D}_c(z) := \text{block-diagonal}\{\mathbf{d}_c(z), \dots, \mathbf{d}_c(z)\}$ with m diagonal block elements all being the same L -column vector $\mathbf{d}_c(z)$ defined by $\mathbf{d}_c(z) = [z^{L_1}, \dots, z^{L_2}]^T$. Therefore, by the similar way to as in [5], the maximization of $|D_{v_i x}|$ under $\sigma_{v_i}^2 = \sigma_{s_{\rho_i}}^2$ leads to the following generalized eigenvector problem;

$$\tilde{\mathbf{B}}\tilde{\mathbf{w}}_i = \lambda_i \tilde{\mathbf{R}}\tilde{\mathbf{w}}_i. \tag{11}$$

Moreover, Jelonnek et al. have shown in [5] that the eigenvector corresponding to the maximum magnitude eigenvalue of $\tilde{\mathbf{R}}^{-1}\tilde{\mathbf{B}}$ becomes the solution of the blind equalization problem, which is referred to as an *eigenvector algorithm* (EVA). It has been also shown in [9] that the BD for MIMO-IIR systems can be achieved

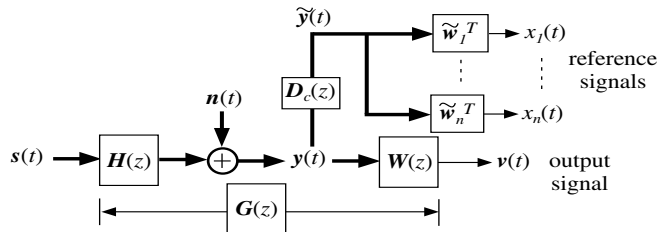


Fig. 2. The composite system of an unknown system and a deconvolver, and a reference system

with the eigenvectors of $\tilde{\mathbf{R}}^\dagger \tilde{\mathbf{B}}$, using only one reference signal. Note that since Jelonnek et al. have dealt with SISO-IIR systems or SIMO-IIR systems, the constructions of $\tilde{\mathbf{B}}$, $\tilde{\mathbf{w}}_i$, and $\tilde{\mathbf{R}}$ in (11) are different from those proposed in [5].

Castella et al. [2] have showed that from (9), a BD can be iteratively achieved by using $x_i(t) = \tilde{\mathbf{w}}_i \tilde{\mathbf{y}}(t)$ ($i = \overline{1, n}$) as reference signals (see Fig. 2), where the number of reference signals corresponds to the number of source signals and $\tilde{\mathbf{w}}_i$ is an eigenvector obtained by $\tilde{\mathbf{R}}^\dagger \tilde{\mathbf{B}}$ in the previous iteration. Then a deflation method was used to recover all source signals.

3.2 The Proposed EVA

In this paper, we want to avoid the conventional EVAs' drawbacks, that is, (a) they are sensitive to Gaussian noise and (b) difference performances are obtained for a different choice of the reference signal $x(t)$. In order to overcome (a), a matrix $\tilde{\mathbf{F}}$ is used instead of $\tilde{\mathbf{R}}$ in (11). This idea comes from [8]. Hence, (11) can be expressed as

$$\tilde{\mathbf{B}} \tilde{\mathbf{w}}_i = \lambda_i \tilde{\mathbf{F}} \tilde{\mathbf{w}}_i, \quad (12)$$

where $\tilde{\mathbf{F}}$ is a set of $m \times m$ block matrices $\mathbf{F}_{\mathbf{y},j,l}^{(4)}$, that is, $\sum_{j=1}^m \sum_{l=L_1}^{L_2} \mathbf{F}_{\mathbf{y},j,l}^{(4)}$, the elements of $\mathbf{F}_{\mathbf{y},j,l}^{(4)}$ are defined by fourth-order cumulants, that is,

$$\begin{aligned} \left[\mathbf{F}_{\mathbf{y},j,l}^{(4)} \right]_{[p,q]l_1 l_2} &= \text{cum}\{y_q(t-L_1-l_2+1), y_p^*(t-L_1-l_1+1), y_j(t-l), y_j^*(t-l)\}, \\ p, q, j &= \overline{1, m}, \quad l_1, l_2 = \overline{1, L}, \quad l = \overline{L_1, L_2}, \end{aligned} \quad (13)$$

Here the matrix $\tilde{\mathbf{B}}$ can be expressed as

$$\tilde{\mathbf{B}} = \tilde{\mathbf{H}}^H \tilde{\mathbf{\Lambda}} \tilde{\mathbf{H}}, \quad (14)$$

where $\tilde{\mathbf{\Lambda}}$ is the block diagonal matrix defined by

$$\tilde{\mathbf{\Lambda}} := \text{block-diag}\{\mathbf{\Lambda}_1, \mathbf{\Lambda}_2, \dots, \mathbf{\Lambda}_n\}, \quad (15)$$

$$\mathbf{\Lambda}_i := \text{diag}\{\dots, |a_i(-1)|^2 \gamma_i, |a_i(0)|^2 \gamma_i, |a_i(1)|^2 \gamma_i, \dots\}, \quad (16)$$

$i = \overline{1, n}$. It is shown by a simple calculation that $\tilde{\mathbf{F}}$ becomes

$$\tilde{\mathbf{F}} = \tilde{\mathbf{H}}^H \tilde{\mathbf{\Psi}} \tilde{\mathbf{H}}, \quad (17)$$

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

$$\tilde{\mathbf{\Psi}} := \text{block-diag}\{\mathbf{\Psi}_1, \mathbf{\Psi}_2, \dots, \mathbf{\Psi}_n\}, \quad (18)$$

$$\mathbf{\Psi}_i := \text{diag}\{\dots, \gamma_i \tilde{a}_i(-1), \gamma_i \tilde{a}_i(0), \gamma_i \tilde{a}_i(1), \dots\}, \quad i = \overline{1, n}, \quad (19)$$

$$\tilde{a}_i(k) := \sum_{j=1}^m \sum_{l=L_1}^{L_2} |h_{ji}(k-l)|^2, \quad i = \overline{1, n}, \quad k \in Z. \quad (20)$$

Let the eigenvalues of the diagonal matrix $\tilde{\mathbf{\Psi}}^{-1} \tilde{\mathbf{\Lambda}}$ be denoted by

$$\mu_i(k) := |a_i(k)|^2 / \tilde{a}_i(k), \quad i = \overline{1, n}, \quad k \in Z. \quad (21)$$

We put the following assumption on the eigenvalues $\mu_i(k)$'s.

A6) All the eigenvalues $\mu_i(k)$'s are distinct for $i = \overline{1, n}$ and $k \in Z$.

Then we can prove the following theorem.

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

T1) All the eigenvalues $\mu_i(k)$'s are distinct for $i = \overline{1, n}$ and $k \in Z$.

Then the n eigenvectors corresponding to n nonzero eigenvalues $\mu_i(k_i)$'s of $\tilde{\mathbf{F}}^\dagger \tilde{\mathbf{B}}$ become the vectors $\tilde{\mathbf{w}}_i$, $i = \overline{1, n}$, satisfying (8),

Outline of the proof: Based on (12), we consider the following eigenvector problem;

$$\tilde{\mathbf{F}}^\dagger \tilde{\mathbf{B}} \tilde{\mathbf{w}}_i = \lambda_i \tilde{\mathbf{w}}_i. \quad (22)$$

Then, from (14) and (17), (22) becomes

$$(\tilde{\mathbf{H}}^H \tilde{\Psi} \tilde{\mathbf{H}})^\dagger \tilde{\mathbf{H}}^H \tilde{\Lambda} \tilde{\mathbf{H}} \tilde{\mathbf{w}}_i = \lambda_i \tilde{\mathbf{w}}_i, \quad (23)$$

Under $L_1 = -\infty$ and $L_2 = \infty$, we have the following equations;

$$(\tilde{\mathbf{H}}^H \tilde{\Psi} \tilde{\mathbf{H}})^\dagger = \tilde{\mathbf{H}}^\dagger \tilde{\Psi}^{-1} \tilde{\mathbf{H}}^{H\dagger}, \quad \tilde{\mathbf{H}}^{H\dagger} \tilde{\mathbf{H}}^H = \mathbf{I}, \quad (24)$$

which are shown in [12] along with their proofs. Then it follows from (23) and (24);

$$\tilde{\mathbf{H}}^\dagger \tilde{\Psi}^{-1} \tilde{\Lambda} \tilde{\mathbf{H}} \tilde{\mathbf{w}}_i = \lambda_i \tilde{\mathbf{w}}_i. \quad (25)$$

Multiplying (25) by $\tilde{\mathbf{H}}$ from the left side and using (24), (25) becomes

$$\tilde{\Psi}^{-1} \tilde{\Lambda} \tilde{\mathbf{H}} \tilde{\mathbf{w}}_i = \lambda_i \tilde{\mathbf{H}} \tilde{\mathbf{w}}_i. \quad (26)$$

$\tilde{\Psi}^{-1} \tilde{\Lambda}$ is a diagonal matrix with diagonal elements $\mu_i(k)$, $i = \overline{1, n}$ and $k \in Z$, and thus (22) and (26) show that its diagonal elements $\mu_i(k)$'s are eigenvalues of matrix $\tilde{\mathbf{F}}^\dagger \tilde{\mathbf{B}}$. Here we use the following fact;

$$\lim_{L \rightarrow \infty} (\text{rank } \tilde{\mathbf{F}}) / L = n, \quad (27)$$

which is shown in [10] and its proof is found in [3]. Using this fact, the other remaining eigenvalues of $\tilde{\mathbf{F}}^\dagger \tilde{\mathbf{B}}$ are all zero. From the assumption **A6)**, the n nonzero eigenvalues $\mu_i(k) \neq 0$, $i = \overline{1, n}$, obtained by (26), that is, the n nonzero eigenvectors $\tilde{\mathbf{w}}_i$, $i = \overline{1, n}$, corresponding to n nonzero eigenvalues $\mu_i(k) \neq 0$, $i = \overline{1, n}$, obtained by (22) become n solutions of the vectors $\tilde{\mathbf{w}}_i$ satisfying (8).

Remark 1. Since the matrix $\tilde{\mathbf{F}}^\dagger \tilde{\mathbf{B}}$ consists of only fourth-order cumulants, the eigenvectors derived from the matrix can be obtained with as little influence of Gaussian noise as possible, which is referred as a *robust eigenvector algorithm* (REVA).

In order to overcome (b), because it is claimed in [2] that the best performance of the EVA can be obtained in the case where the recovered signals are used as the reference signals, we use $x_i(t) = \tilde{\mathbf{w}}_i \tilde{\mathbf{y}}(t)$ in Fig. 2 as the reference signals,

where the number of the reference signals corresponds to the number of the source signals. However, the EVA using the reference signals needs deflation methods to recover all source signals (see [2]). On the contrary, we don't want to use deflation methods because the deflation is affected to the estimation errors and hence as the deflation process comes near to the final step, the accuracy of recovering source signals is getting worse. Therefore, the following two-stage EVA is proposed.

Stage 1) Roughly estimate the eigenvectors of $\tilde{\mathbf{R}}^\dagger \tilde{\mathbf{B}}$ with the reference signal $x(t)$ in (9)

In Stage 1), all the vectors $\tilde{\mathbf{w}}_i$ corresponding to (8) can be simultaneously obtained using only one reference signal $x(t)$ (see [9]). Since the estimate of $\tilde{\mathbf{R}}$ has a good accuracy with a few samples, compared with the estimate of $\tilde{\mathbf{F}}$, then first of all the eigenvectors are roughly estimated with $\tilde{\mathbf{R}}^\dagger \tilde{\mathbf{B}}$. However, the vectors obtained by $\tilde{\mathbf{R}}^\dagger \tilde{\mathbf{B}}$ are sensitive to Gaussian noise and their performances depend on reference signals (see Section 4). Then,

Stage 2) Estimate the eigenvectors of $\tilde{\mathbf{F}}^\dagger \tilde{\mathbf{B}}_i$ with the reference signals $x_i(t)$ $\tilde{\mathbf{w}}_i \tilde{\mathbf{y}}(t)$, $i = \overline{1, n}$,

where $\tilde{\mathbf{w}}_i$ ($i = \overline{1, n}$) are the eigenvectors obtained in Stage 1) and the matrix $\tilde{\mathbf{B}}$ obtained by using each $x_i(t)$ is denoted by $\tilde{\mathbf{B}}_i$.

Each eigenvector obtained in Stage 2) is the one corresponding to the absolute maximum eigenvalue $|\lambda_i|$ for each $\tilde{\mathbf{F}}^\dagger(t) \tilde{\mathbf{B}}_i(t)$. Although $x_i(t) = \tilde{\mathbf{w}}_i \tilde{\mathbf{y}}(t)$ ($i = \overline{1, n}$) are used as the reference signals (see Fig. 2), deflation methods are not needed to recover all source signals in this stage, because the vectors $\tilde{\mathbf{w}}_i$'s in $x_i(t)$'s, which are obtained in Stage 1), have been already linearly independent.

If the matrices of $\tilde{\mathbf{R}}$, $\tilde{\mathbf{F}}$, and $\tilde{\mathbf{B}}$ can be estimated with good accuracies, the two-stage algorithm can provide the solution in (8) with one iteration. In this paper, however, since we confine ourselves to implement their estimates iteratively, the procedure of the two-stage EVA is summarized as follows:

Choose an appropriate reference signal $x(t)$ and appropriate initial values of

$\tilde{\mathbf{w}}_i^{[1]}(0)$, $\tilde{\mathbf{w}}_i^{[2]}(0)$, $\tilde{\mathbf{R}}(0)$, $\tilde{\mathbf{F}}(0)$, $\tilde{\mathbf{B}}(0)$, $\tilde{\mathbf{B}}_i(0)$

for $t_l = 1 : t_{all}$

if $t_l < t_s$

for $t = t_d(t_l - 1) + 1 : t_d t_l$

Calculate $\tilde{\mathbf{R}}(t)$, $\tilde{\mathbf{F}}(t)$, and $\tilde{\mathbf{B}}(t)$ by their moving averages.

end

Calculate the eigenvectors $\tilde{\mathbf{w}}_i^{[1]}(t_l)$'s from $\tilde{\mathbf{R}}^\dagger(t) \tilde{\mathbf{B}}(t)$ (Stage 1)).

elseif $t_l \geq t_s$

if $t_l == t_s$, $\tilde{\mathbf{w}}_i^{[2]}(t_l - 1) = \tilde{\mathbf{w}}_i^{[1]}(t_s - 1)$

for $t = t_d(t_l - 1) + 1 : t_d t_l$

$x_i(t) = \tilde{\mathbf{w}}_i^{[2]T}(t_l - 1) \tilde{\mathbf{y}}_i(t)$

Calculate $\tilde{\mathbf{F}}(t)$ and $\tilde{\mathbf{B}}_i(t)$ by their moving averages.

end

Calculate the eigenvector $\tilde{\mathbf{w}}_i^{[2]}(t_l)$ associated with the absolute maximum eigenvalue $|\lambda_i|$ from $\tilde{\mathbf{F}}^\dagger(t)\tilde{\mathbf{B}}_i(t)$, $i = \overline{1, n}$ (Stage 2).
end
end

Here, $t_{l_{all}}$ denotes the total number of iterations and t_d denotes the number of data samples for estimating the matrices $\tilde{\mathbf{R}}(t)$, $\tilde{\mathbf{F}}(t)$, $\tilde{\mathbf{B}}(t)$, and $\tilde{\mathbf{B}}_i(t)$. From a practical viewpoint, it would be better to estimate the fourth-order cumulant matrix $\tilde{\mathbf{F}}(t)$ during $t_l = 1$ to $t_{l_{all}}$. $\tilde{\mathbf{w}}_i^{[1]}(t)$ and $\tilde{\mathbf{w}}_i^{[2]}(t)$ are the eigenvectors obtained for Stage 1) and Stage 2), respectively. t_s denotes an arbitrary integer satisfying $2 < t_s < t_{l_{all}}$. For $0 < t_l < t_s$ and $t_s \leq t_l \leq t_{l_{all}}$, the eigenvectors $\tilde{\mathbf{w}}_i^{[1]}(t_l)$'s and $\tilde{\mathbf{w}}_i^{[2]}(t_l)$'s are iteratively calculated, respectively, according to Stage 1) and Stage 2).

4 Computer Simulations

To demonstrate the validity of the proposed algorithm, many computer simulations were conducted. Some results are shown in this section. The unknown system $\mathbf{H}(z)$ was set to be the same system with two inputs and three outputs as in [9]. The source signals $s_1(t)$ and $s_2(t)$ were a sub-Gaussian signal which takes one of two values, -1 and 1 with equal probability $1/2$. The Gaussian noises $n_j(t)$ with its variance $\sigma_{n_j}^2$ were included in the output $y_j(t)$ at various SNR levels. The SNR was considered at the output of the system $\mathbf{H}(z)$. The parameters L_1 and L_2 in $\mathbf{W}(z)$ were set to be 0 and 11, respectively. As a measure of performances, we used the *multichannel intersymbol interference* (M_{ISI}) [11], which was the average of 50 Monte Carlo runs. In each Monte Carlo run, the number of iterations $t_{l_{all}}$ was set to be 10, the number of data samples t_d was set to be 5,000, and the threshold t_s was set to be 6.

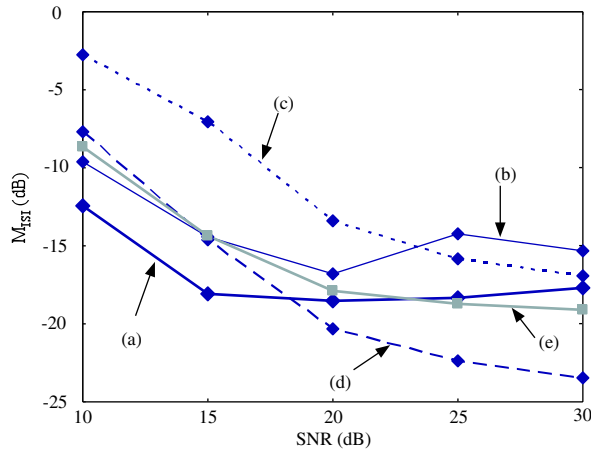


Fig. 3. The performances of the proposed algorithm and the conventional EVAs with varying SNR levels, in the cases of 5,000 data samples

Fig. 3 shows the results of performances of the EVAs when the SNR levels were respectively taken to be 10 through 30 dB for every 5 dB. The line (a) represents the performance obtained by our proposed two-stage EVA, where in the EVA in the stage 1), $x(t) = f_2(2)y_2(t-2)$ was used as the reference signal and $f_2(2)$ was randomly chosen from a Gaussian distribution with zero mean and unit variance. The line (b) represents the performance obtained by only the REVA ($t_s = 1$), where the reference signal was given by the same equation as the line (a), but the parameter $f_2(2)$ was given by another Gaussian distribution with zero mean and unit variance. The lines (c) and (d) represent the performances obtained by only the EVA ($t_s = 11$) with respectively $x(t) = \sum_{i=1}^3 f_i(2)y_i(t-2)$ and $f_2(2)y_2(t-2)$ as the reference signal. Finally, the line (e) represents the performance obtained by Castella et al. (CRMPA). From the line (b), the REVA has such a property that as the SNR level decreases, it can provide better performances than the EVA. That is, the REVA can work with as little influence of Gaussian noise as possible. However for the high SNR levels its performance is not so good compared with the EVA. From the lines (c) and (d), one can see that the EVA's performance depends on the choice of the reference signal. From the line (a), our proposed algorithm has such a property that as the SNR level decreases, it is more robust to Gaussian noise than the other algorithms, and for the high SNR levels, its performance is almost the same as the CRMPA. Therefore from all the results we conclude that our proposed algorithm can overcome the drawbacks of the conventional EVAs.

5 Conclusions

We have proposed a two-stage algorithm obtained by combining the EVA and the REVA for solving the BD problem. The proposed algorithm provides better performances than the conventional EVAs, because the proposed algorithm can overcome the drawbacks of the conventional EVAs, that is,

- The EVA is sensitive to Gaussian noise.
- The EVA depends on the selection of the reference signal.

Computer simulations have demonstrated the effectiveness of the proposed EVA.

Acknowledgments. This work is supported by the Research Projects, Grant-IN AIDs, No. 18500146¹ and No. 1850054² for Scientific Research of the JSPS.

References

1. Adib, A., et al.: Source separation contrasts using a reference signal. *IEEE Signal Processing Letters* 11(3), 312–315 (2004)
2. Castella, M., et al.: Quadratic Higher-Order Criteria for Iterative Blind Separation of a MIMO Convolutional Mixture of Sources. *IEEE Trans. Signal Processing* 55(1), 218–232 (2007)

3. Inouye, Y.: Autoregressive model fitting for multichannel time series of degenerate rank: Limit properties. *IEEE Trans. Circuits and Systems* 32(3), 252–259 (1985)
4. Inouye, Y., Tanebe, K.: Super-exponential algorithms for multichannel blind deconvolution. *IEEE Trans. Sig. Proc.* 48(3), 881–888 (2000)
5. Jelonnek, B., Kammeyer, K.D.: A closed-form solution to blind equalization. *Signal Processing* 36(3), 251–259 (1994)
6. Jelonnek, B., Boss, D., Kammeyer, K.D.: Generalized eigenvector algorithm for blind equalization. *Signal Processing* 61(3), 237–264 (1997)
7. Kawamoto, M., et al.: Eigenvector algorithms using reference signals. In: *Proc. ICASSP 2006, May 2006, vol. V*, pp. 841–844 (2006)
8. Kawamoto, M., et al.: Eigenvector algorithms using reference signals for blind source separation of instantaneous mixtures. In: *Proc. of ISCAS 2006, May 2006*, pp. 4191–4194 (2006)
9. Kawamoto, M., et al.: Eigenvector algorithms for blind deconvolution of MIMO-IIR systems. In: *Proc. ISCAS 2007, May 2007*, pp. 3490–3493 (2007)
10. Kohno, K., et al.: Adaptive super-exponential algorithms for blind deconvolution of MIMO systems. In: *Proc. ISCAS 2004, May 2004, vol. V*, pp. 680–683 (2004)
11. Kohno, K., et al.: Super-Exponential Methods Incorporated with Higher-Order Correlations for Deflationary Blind Equalization of MIMO Linear Systems. In: *Proc. ICA 2004*, pp. 685–693 (2004)
12. Kohno, K., et al.: Robust super-exponential methods for blind equalization of MIMO-IIR systems. In: *Proc. ICASSP 2006, vol. V*, pp. 661–664 (2006)
13. Parra, L., Sajda, P.: Blind source separation via generalized eigenvalue decomposition. *Journal of Machine Learning* (4), 1261–1269 (2003)
14. Rhioui, S., et al.: Quadratic MIMO contrast functions for blind source separation in a convolutive context. In: *Proc. ICA 2006*, pp. 230–237 (2006)