Eigenvector Algorithms for Blind Deconvolution of MIMO-IIR Systems

Mitsuru Kawamoto[1]
National Institute of Advanced Industrial Science and Technology (AIST),
Central 2, 1-1-1 Umezono,
Tsukuba, Ibaraki 305–8568, Japan
Email: m.kawamoto@aist.go.jp

Kiyotaka Kohno* and Yujiro Inouye†2
Department of Electronic and Control Systems Engineering,
Shimane University
1060 Nishikawatsu, Matsue, 690-8504, Japan
Email: *kohno@yonago-k.ac.jp
Email: †inouye@riko.shimane-u.ac.jp

***Abstract*— This paper presents eigenvector algorithms (EVAs) for blind deconvolution (BD) of multiple-input multiple-output infinite impulse response (MIMO-IIR) channels (convolutive mixtures). Using the idea of reference signals, the EVA is derived. Differently from the conventional researches on EVAs, one of the novel points of the paper is that the EVA using *any reference signal* is applied to the BD problem of the MIMO-IIR system, and then the validity of the EVA is shown.**

I. Introduction

In this paper, we deal with a blind deconvolution (BD) problem for a multiple-input and multiple-output (MIMO) infinite-impulse response (IIR) channels. To solve this problem, we use eigenvector algorithms (EVAs) [4], [8]. The first proposal of the EVA was done by Jelonnek et al. [4]. They have proposed the EVA for solving blind equalization (BE) problems of single-input single-output (SISO) channels or single-input multiple-output (SIMO) channels. In [8], several procedures for the blind source separation (BSS) of instantaneous mixtures, using the generalized eigenvalue decomposition, have been introduced. Recently, the authors have proposed an EVA which can solve blind source separation (BSS) problems in the case of MIMO static systems (instantaneous mixtures) [5].

The proposed EVA, based on the idea in [5], is derived by using reference signals. Researches using the idea of reference signals to solve blind signal processing (BSP) problems, such as the BD, the BE, the BSS, and so on, to our best knowledge, have been made by Jelonnek et al. (e.g., [4]), Adib et al. (e.g., [1]), and Rhioui [9]. 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 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. 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 which can be extracted. Moreover, they claimed that as the reference signal, it is a practical valid choice to choose a signal obtained by whitening the outputs of the MIMO convolved system.

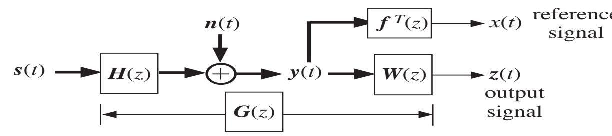

Fig. 1. The composite system of an unknown system and a deconvolver, and a reference system.

In this paper, under the assumption that *any reference signal* is used, we want to show how the EVA works for the BD of the MIMO-IIR channel (1). Simulation results are presented to show the effectiveness of the proposed EVA.

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 $\boldsymbol{C}^n$ 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, $*$, and H denote, respectively, the transpose, the complex conjugate, and the complex conjugate transpose (Hermitian) of a matrix. The symbols block-diag$\{\cdots\}$ and diag$\{\cdots\}$ denote respectively a block diagonal and a diagonal matrices 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 $i = \overline{1,n}$ stands for $i = 1, 2, \cdots, n$.

II. Problem Formulation and Assumptions

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

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

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 channel is defined by $\boldsymbol{H}(z) = \sum_{k=-\infty}^{\infty} \boldsymbol{H}^{(k)} z^k, z \in C$.

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

$$\begin{aligned}\boldsymbol{z}(t) &= \sum_{k=-\infty}^{\infty} \boldsymbol{W}^{(k)} \boldsymbol{y}(t-k) \\ &= \sum_{k=-\infty}^{\infty} \boldsymbol{G}^{(k)} \boldsymbol{s}(t-k) + \sum_{k=-\infty}^{\infty} \boldsymbol{W}^{(k)} \boldsymbol{n}(t-k),\end{aligned} \tag{2}$$

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

Here, we put the following assumptions on the 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 $\sigma^2_{s_i} \neq 0$ and a nonzero fourth-order cumulant $\gamma_i \neq 0$ defined as

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

A3) The deconvolver $\boldsymbol{W}(z)$ is an 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 $\sigma^2_{n_j}$, $j = \overline{1,m}$.

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

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

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

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

$$\tilde{\boldsymbol{g}}_i = \tilde{\boldsymbol{H}}\tilde{\boldsymbol{w}}_i, \quad i = \overline{1,n}, \tag{5}$$

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

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

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

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

$j = \overline{1,m}$, 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 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 $[\boldsymbol{H}_{ij}]_{lr} := h_{ji}(l-r)$, $l = 0, \pm1, \pm2, \cdots$, $r = \overline{L_1, L_2}$, 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

$$[\tilde{\boldsymbol{g}}_1, \cdots, \tilde{\boldsymbol{g}}_n] = \tilde{\boldsymbol{H}}[\tilde{\boldsymbol{w}}_1, \cdots, \tilde{\boldsymbol{w}}_n] = [\tilde{\boldsymbol{\delta}}_1, \cdots, \tilde{\boldsymbol{\delta}}_n]\boldsymbol{P}, \tag{8}$$

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

$$\tilde{\boldsymbol{\delta}}_i := [\boldsymbol{\delta}_{i1}^T, \boldsymbol{\delta}_{i2}^T, \ldots, \boldsymbol{\delta}_{in}^T]^T, \quad i = \overline{1,n} \tag{9}$$

$$\boldsymbol{\delta}_{ij} := \begin{cases} \hat{\boldsymbol{\delta}}_i, & \text{if} \quad i = j, \\ (\cdots, 0, 0, 0, \cdots)^T, & \text{otherwise.} \end{cases} \tag{10}$$

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

III. Eigenvector algorithms (EVAs)

A. Analysis of eigenvector algorithms with reference signals for MIMO-IIR channels

In order to solve the BD problem, the following cross-cumulant between $z_i(t)$ and a reference signal $x(t)$ (see Fig. 1) is defined;

$$D_{zx} = \mathrm{cum}\{z_i(t), z_i^*(t), x(t), x^*(t)\}, \tag{11}$$

where $z_i(t)$ is the ith element of $\boldsymbol{z}(t)$ in (2) and the reference signal $x(t)$ is given by $\boldsymbol{f}^T(z)\boldsymbol{y}(t)$, using an appropriate filter $\boldsymbol{f}(z)$. The filter $\boldsymbol{f}(z)$ is called a *reference system*. Let $\boldsymbol{a}(z) := \boldsymbol{H}^T(z)\boldsymbol{f}(z) = [a_1(z), a_2(z), \cdots, a_n(z)]^T$, 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 $a_i(z) = \sum_{k=-\infty}^{\infty} a_i(k)z^k$ and the reference system $\boldsymbol{f}(z)$ is an m-column vector whose elements are $f_j(z) = \sum_{k=L_1}^{L_2} f_j(k)z^k$, $j = \overline{1,m}$.

Jelonnek et al. [4] have shown in the single-input case that by the Lagrangian method, the maximization of $|D_{zx}|$ under $\sigma^2_{z_i} = \sigma^2_{s_{\rho_i}}$ leads to a closed-form expressed as a generalized eigenvector problem, where $\sigma^2_{z_i}$ and $\sigma^2_{s_{\rho_i}}$ denote the variances of the output $z_i(t)$ and a source signal $s_{\rho_i}(t)$, respectively, and ρ_i is one of integers $\{1, 2, \cdots, n\}$ such that the set $\{\rho_1, \rho_2, \cdots, \rho_n\}$ is a permutation of the set $\{1, 2, \cdots, n\}$. In our case, D_{zx} and $\sigma^2_{z_i}$ can be expressed in terms of the vector $\tilde{\boldsymbol{w}}_i$ as, respectively,

$$D_{zx} = \tilde{\boldsymbol{w}}_i^H \tilde{\boldsymbol{B}} \tilde{\boldsymbol{w}}_i, \;\; \sigma^2_{z_i} = \tilde{\boldsymbol{w}}_i^H \tilde{\boldsymbol{R}} \tilde{\boldsymbol{w}}_i, \tag{12}$$

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

$$\tilde{\boldsymbol{y}}(t) := \left[\boldsymbol{y}_1^T(t), \boldsymbol{y}_2^T(t), \cdots, \boldsymbol{y}_m^T(t)\right]^T \in \boldsymbol{C}^{mL}, \tag{13}$$

$$\boldsymbol{y}_j(t) := [y_j(t\text{-}L_1), y_j(t\text{-}L_1\text{-}1), \cdots, y_j(t\text{-}L_2)]^T \in \boldsymbol{C}^L, \tag{14}$$

$j = \overline{1,m}$. Therefore, by the similar way to as in [4], the maximization of $|D_{zx}|$ under $\sigma^2_{z_i} = \sigma^2_{s_{\rho_i}}$ leads to the following generalized eigenvector problem;

$$\tilde{\boldsymbol{B}}\tilde{\boldsymbol{w}}_i = \lambda_i \tilde{\boldsymbol{R}}\tilde{\boldsymbol{w}}_i. \tag{15}$$

Moreover, Jelonnek et al. have shown that the eigenvector corresponding to the maximum magnitude eigenvalue of $\tilde{\boldsymbol{R}}^{\dagger}\tilde{\boldsymbol{B}}$ becomes the solution of the blind equalization problem in [4], 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 (15) are different from those proposed in [4]. In this paper, under the assumption that any reference system $\boldsymbol{f}(z)$ is used, we want to show how the eigenvector algorithm (15) works for the BD of the MIMO-IIR channel (1).

To this end, we use the following equalities;

$$\tilde{\boldsymbol{R}} = \tilde{\boldsymbol{H}}^H \tilde{\boldsymbol{\Sigma}} \tilde{\boldsymbol{H}}, \;\; \tilde{\boldsymbol{B}} = \tilde{\boldsymbol{H}}^H \tilde{\boldsymbol{\Lambda}} \tilde{\boldsymbol{H}}, \tag{16}$$

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

$$\tilde{\boldsymbol{\Sigma}} := \text{block-diag}\{\boldsymbol{\Sigma}_1, \boldsymbol{\Sigma}_2, \cdots, \boldsymbol{\Sigma}_n\}, \quad (17)$$

$$\boldsymbol{\Sigma}_i := \text{diag}\{\cdots, \sigma_{s_i}^2, \sigma_{s_i}^2, \sigma_{s_i}^2, \cdots\}, \quad i = \overline{1,n}, \quad (18)$$

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

$$\tilde{\boldsymbol{\Lambda}} := \text{block-diag}\{\boldsymbol{\Lambda}_1, \boldsymbol{\Lambda}_2, \cdots, \boldsymbol{\Lambda}_n\}, \quad (19)$$

$$\boldsymbol{\Lambda}_i := \text{diag}\{\cdots, |a_i(-1)|^2\gamma_r, |a_i(0)|^2\gamma_i, |a_i(1)|^2\gamma_i, \cdots\}, \quad (20)$$

$i = \overline{1,n}$. Since both $\tilde{\boldsymbol{\Sigma}}$ and $\tilde{\boldsymbol{\Lambda}}$ become diagonal, (16) shows that the two matrices $\tilde{\boldsymbol{R}}$ and $\tilde{\boldsymbol{B}}$ are simultaneously diagonalizable in a wide sense.

Here, let the eigenvalues of the diagonal matrix $\tilde{\boldsymbol{\Sigma}}^{-1}\tilde{\boldsymbol{\Lambda}}$ is denoted by

$$\lambda_i(k) := |a_i(k)|^2\gamma_i/\sigma_{s_i}^2, \quad i = \overline{1,n}, \quad k \in Z. \quad (21)$$

We put the following assumption on the eigenvalues $\lambda_i(k)'s$.
A6) All the eigenvalues $\lambda_i(k)'s$ are distinct for $i = \overline{1,n}$ and $k \in Z$.

Theorem 1: Suppose the noise term $\boldsymbol{n}(t)$ is absent and the length L of the deconvolver is infinite (that is, $L_1 = -\infty$ and $L_2 = \infty$). Then, under the assumptions **A1)** through **A6)**, the n eigenvector $\tilde{\boldsymbol{w}}_i$'s corresponding to the n nonzero eigenvalues $\lambda_i(k)'s$ of matrix $\tilde{\boldsymbol{R}}^\dagger\tilde{\boldsymbol{B}}$ for $i = \overline{1,n}$ and an arbitrary $k \in Z$ become the vectors $\tilde{\boldsymbol{w}}_i$'s satisfying (8).

Outline of the proof: Based on (15), we consider the following eigenvector problem;

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

Then, from (16), (22) becomes

$$(\tilde{\boldsymbol{H}}^H\tilde{\boldsymbol{\Sigma}}\tilde{\boldsymbol{H}})^\dagger\tilde{\boldsymbol{H}}^H\tilde{\boldsymbol{\Lambda}}\tilde{\boldsymbol{H}}\tilde{\boldsymbol{w}}_i = \lambda_i\tilde{\boldsymbol{w}}_i. \quad (23)$$

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

$$(\tilde{\boldsymbol{H}}^H\tilde{\boldsymbol{\Sigma}}\tilde{\boldsymbol{H}})^\dagger = \tilde{\boldsymbol{H}}^\dagger\tilde{\boldsymbol{\Sigma}}^\dagger\tilde{\boldsymbol{H}}^{H\dagger}, \;\; \tilde{\boldsymbol{H}}^{H\dagger}\tilde{\boldsymbol{H}}^H = \boldsymbol{I}, \quad (24)$$

which are shown in [7] along with their proofs. Then it follows from (23) and (24);

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

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

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

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

$$\lim_{L\to\infty}(\text{rank } \tilde{\boldsymbol{R}})/L = n, \quad (27)$$

which is shown in [6] and its proof is found in [2]. Using this fact, the other remaining eigenvalues of $\tilde{\boldsymbol{R}}^\dagger\tilde{\boldsymbol{B}}$ are all zero. From the assumption **A6)**, the n nonzero eigenvalues $\lambda_i(k) \neq 0$, $i = \overline{1,n}$, obtained by (26), that is, the n nonzero eigenvectors $\tilde{\boldsymbol{w}}_i$, $i = \overline{1,n}$, corresponding to n nonzero eigenvalues $\lambda_i(k) \neq 0$, $i = \overline{1,n}$, obtained by (22) become n solutions of the vectors $\tilde{\boldsymbol{w}}_i$ satisfying (8).

Remark 1: When the length L of the deconvolver is finite, the size of the matrix $\tilde{\boldsymbol{R}}^\dagger\tilde{\boldsymbol{B}}$ is $mL \times mL$, but its rank is asymptotically equal to nL as $L \to \infty$. Therefore, it follows from the assumption **A6)** that there exist nL nonzero eigenvalues of $\tilde{\boldsymbol{R}}^\dagger\tilde{\boldsymbol{B}}$ which are approximately equal to the n nonzero eigenvalues $\lambda_i(k)$, $i = \overline{1,n}$ of the matrix $\tilde{\boldsymbol{\Sigma}}^{-1}\tilde{\boldsymbol{\Lambda}}$ and $(m-n)L$ eigenvalues of $\tilde{\boldsymbol{R}}^\dagger\tilde{\boldsymbol{B}}$ which are approximately equal to zero.

B. How to choose the eigenvectors

From Remark 1, we have a problem of how the eigenvectors corresponding to $\tilde{\boldsymbol{w}}_i$, $i = \overline{1,n}$, satisfying (8) can be chosen from all eigenvectors of $\tilde{\boldsymbol{R}}^\dagger\tilde{\boldsymbol{B}}$. In this subsection, a solution of the problem will be shown. To this end, we consider the following eigenvector problem;

$$\tilde{\boldsymbol{B}}\tilde{\boldsymbol{R}}^\dagger\hat{\boldsymbol{w}}_i = \hat{\lambda}_i\hat{\boldsymbol{w}}_i, \quad (28)$$

where the structure of $\hat{\boldsymbol{w}}_i$ is the same as the one of $\tilde{\boldsymbol{w}}_i$, but the elements of $\hat{\boldsymbol{w}}_i$ are different from the ones of $\tilde{\boldsymbol{w}}_i$. The eigenvalues $\hat{\lambda}_i$'s of $\tilde{\boldsymbol{B}}\tilde{\boldsymbol{R}}^\dagger$ correspond to λ_i's of $\tilde{\boldsymbol{R}}^\dagger\tilde{\boldsymbol{B}}$, because the eigenvectors obtained from (28) are the left eigenvectors of $\tilde{\boldsymbol{R}}^\dagger\tilde{\boldsymbol{B}}$, corresponding to λ_i's. Moreover, the conjugately transposed vectors of the eigenvectors obtained from (28) correspond (or are equal) to the row vectors of $\tilde{\boldsymbol{H}}$ in (5) up to constants. The proof of the mentioned above is given below: Substituting (16) into (28), we obtain

$$\tilde{\boldsymbol{H}}^H\tilde{\boldsymbol{\Lambda}}\tilde{\boldsymbol{H}}(\tilde{\boldsymbol{H}}^H\tilde{\boldsymbol{\Sigma}}\tilde{\boldsymbol{H}})^\dagger\hat{\boldsymbol{w}}_i = \hat{\lambda}_i\hat{\boldsymbol{w}}_i. \quad (29)$$

By the similar way to (25), (29) becomes

$$\tilde{\boldsymbol{H}}^H\tilde{\boldsymbol{\Lambda}}\tilde{\boldsymbol{\Sigma}}^{-1}\tilde{\boldsymbol{H}}^{H\dagger}\hat{\boldsymbol{w}}_i = \hat{\lambda}_i\hat{\boldsymbol{w}}_i, \quad (30)$$

Multiplying (30) by $\tilde{\boldsymbol{H}}^{H\dagger}$ from the left side, (30) becomes

$$\tilde{\boldsymbol{\Lambda}}\tilde{\boldsymbol{\Sigma}}^{-1}\tilde{\boldsymbol{H}}^{H\dagger}\hat{\boldsymbol{w}}_i = \hat{\lambda}_i\tilde{\boldsymbol{H}}^{H\dagger}\hat{\boldsymbol{w}}_i. \quad (31)$$

Let $\hat{\boldsymbol{g}}_i := \tilde{\boldsymbol{H}}^{H\dagger}\hat{\boldsymbol{w}}_i$, then (31) becomes $\tilde{\boldsymbol{\Lambda}}\tilde{\boldsymbol{\Sigma}}^{-1}\hat{\boldsymbol{g}}_i = \hat{\lambda}_i\hat{\boldsymbol{g}}_i$. This means that since $\tilde{\boldsymbol{\Lambda}}\tilde{\boldsymbol{\Sigma}}^{-1}$ is a diagonal matrix, the elements of $\hat{\boldsymbol{g}}_i$ are zero except for one element. On the other hand, multiplying $\hat{\boldsymbol{g}}_i = \tilde{\boldsymbol{H}}^{H\dagger}\hat{\boldsymbol{w}}_i$ by $\tilde{\boldsymbol{H}}^H$ from the left side, we have

$$\tilde{\boldsymbol{H}}^H\hat{\boldsymbol{g}}_i = \tilde{\boldsymbol{H}}^H\tilde{\boldsymbol{H}}^{H\dagger}\hat{\boldsymbol{w}}_i. \quad (32)$$

We obtain from (29) that $\hat{\boldsymbol{w}}_i$ belongs to the range of $\tilde{\boldsymbol{H}}^H$. This fact means that there exists a vector $\hat{\boldsymbol{\xi}}$ such that $\hat{\boldsymbol{w}}_i = \tilde{\boldsymbol{H}}^H\hat{\boldsymbol{\xi}}_i$. Since $\tilde{\boldsymbol{H}}^H\tilde{\boldsymbol{H}}^{H\dagger}\tilde{\boldsymbol{H}}^H = \tilde{\boldsymbol{H}}^H$, (32) gives

$$\tilde{\boldsymbol{H}}^H\hat{\boldsymbol{g}}_i = \tilde{\boldsymbol{H}}^H\tilde{\boldsymbol{H}}^{H\dagger}\hat{\boldsymbol{w}}_i = \tilde{\boldsymbol{H}}^H\tilde{\boldsymbol{H}}^{H\dagger}\tilde{\boldsymbol{H}}^H\hat{\boldsymbol{\xi}}_i = \tilde{\boldsymbol{H}}^H\hat{\boldsymbol{\xi}}_i = \hat{\boldsymbol{w}}_i. \quad (33)$$

which implies $\hat{\boldsymbol{w}}_i^H = \hat{\boldsymbol{g}}_i^H\tilde{\boldsymbol{H}}$. This shows along with the fact that all the elements of $\hat{\boldsymbol{g}}_i$ are zero except for one element that the conjugately transposed vector of $\hat{\boldsymbol{w}}_i$ becomes a row vector of $\tilde{\boldsymbol{H}}$ up to a constant. This completes the proof.

It can be seen from the definition of the block element $\boldsymbol{H}_{ij}$ (see it stated below (7)) that $\boldsymbol{H}_{ij}$ is a matrix (of L columns and possibly infinite number of rows) having a special Toeplitz (or constant-along-diagonals) structure. Therefore, the (cross) correlation of a pair of rows of $\boldsymbol{H}_{ij}$ (by shifting their elements left or right appropriately) is the same for all pairs of rows of $\boldsymbol{H}_{ij}$ if L is infinite. In practice, however, the length L of the equalizer and the length K of the channel are finite, and so $\boldsymbol{H}_{ij}$ is a matrix of L columns and $L + K - 1$ rows, that is, $\boldsymbol{H}_{ij} \in \boldsymbol{C}^{(L+K-1)\times L}$. In this case, pairs of rows of $\boldsymbol{H}_{ij}$ have approximately the similar correlations for all pairs of rows of $\boldsymbol{H}_{ij}$ if L is sufficiently large. According to Remark 1 and the above discussion, we consider nL nonzero eigenvalues and $(m-n)L$ approximately-zero eigenvalues of the matrix $\tilde{\boldsymbol{R}}^{\dagger}\tilde{\boldsymbol{B}}$, and we can classify approximately nL eigenvectors $\tilde{\boldsymbol{w}}_i$ in (22) corresponding to nL nonzero eigenvalues into n sets of L eigenvectors whose pair have almost the same correlations for all pair of eigenvectors of each set, and there remain $(m-n)L$ eigenvectors corresponding to the remaining $(m-n)L$ eigenvalues which are approximately zero. Thus we propose a tentative procedure of finding n appropriate eigenvectors satisfying (8) is as follows;

1) Set k = 1 (where k denotes the number of iterations from the beginning less than n+1).

2) Select the eigenvector $\tilde{\boldsymbol{w}}$ of $\tilde{\boldsymbol{R}}^{-1}\tilde{\boldsymbol{B}}$ and the eigenvector $\hat{\boldsymbol{w}}$ of $\tilde{\boldsymbol{B}}\tilde{\boldsymbol{R}}^{-1}$ corresponding to the maximum magnitude eigenvalue among $|\hat{\lambda}_i|$'s.

3) Calculate the magnitude of the correlations of all pairs of $\hat{\boldsymbol{w}}$ and $\hat{\boldsymbol{w}}_i$'s.

4) Separate L eigenvalues $\hat{\lambda}$'s from the others such that their magnitudes are larger than the remaining $(n-k)L$ ones, and save the $(n-k)L$ remaining eigenvalues $\hat{\lambda}$'s for finding other eigenvectors.

5) Put k = k + 1 and stock the $\tilde{\boldsymbol{w}}$ obtained in **2)**. If k = n + 1, stop the iterations, otherwise, go to **2)**.

Therefore, the n eigenvectors $\tilde{\boldsymbol{w}}$'s stocked in step **5)** are the n solutions in (8)

IV. 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 three outputs, and assumed that the length of channel was three ($K=3$), that is, $\boldsymbol{H}^{(k)}$'s in (1) were set to be $\boldsymbol{H}(z) = \sum_{k=0}^{2}\boldsymbol{H}^{(k)}z^k =$

$$\begin{bmatrix} 1.00-0.35z+0.10z^2 & 0.65+0.25z-0.15z^2 \\ 0.50-0.30z+0.20z^2 & 1.00+0.25z+0.10z^2 \\ 0.60+0.10z+0.40z^2 & 0.10+0.20z+0.10z^2 \end{bmatrix}.$$

The Gaussian noises $n_j(t)$ with its variance $\sigma^2_{n_j}$ were included in the output $y_j(t)$ at various SNR levels. The SNR was considered at the output of the system $\boldsymbol{H}(z)$. 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. As a measure of performances, we used the $multichannel\ intersymbol\ interference$ ($\mathrm{M_{ISI}}$) [3]. The parameters L_1 and L_2 in $\boldsymbol{W}(z)$ were set to be 0 and 9, respectively. The first and the second components of the reference system $\boldsymbol{f}(z)$ were, respectively, set to be z^2 and 0. that is, $x(t) = y_1(t-2)$.

Fig. 2 shows the results of performances of the proposed EVA when the SNR levels were respectively taken to be 5 through 40 dB for every 5 dB, where each $\mathrm{M_{ISI}}$ shown in Fig. 2 was the average of the performances obtained by 30 independent Monte Carlo runs. In each Monte Carlo run, the final eigenvectors were obtained by ten iterative calculations, where in each iteration, $\tilde{\boldsymbol{R}}$ and $\tilde{\boldsymbol{B}}$ were estimated by 5,000 data samples.

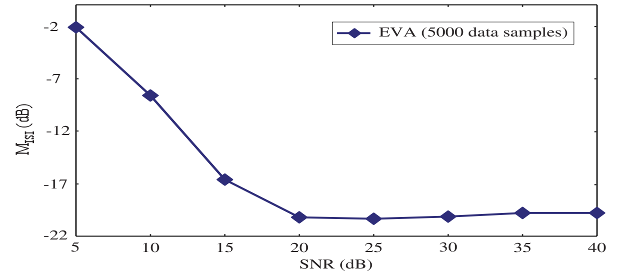

Fig. 2. The performances of the EVA with varying SNR levels, in the cases of 5,000 data samples.

It can be seen from Fig. 2 that when the SNR level is more than about 15 dB, the EVA provides good performances. However, as the SNR level decreases, the performances obtained by the EVA become worse. The reason is that since the matrix $\tilde{\boldsymbol{R}}$ of $\tilde{\boldsymbol{R}}^{\dagger}\tilde{\boldsymbol{B}}$ is a covariance matrix of $\tilde{\boldsymbol{y}}(t)$ in (13), that is, $\tilde{\boldsymbol{R}}$ is formulated by second-order statistics, then the eigenvector calculation is affected by the Gaussian noise.

V. Conclusions

We have proposed an EVA for solving the BD problem. Differently from the methods which needs deflation methods, e.g., super-exponential methods, the EVA can provide such filters that source signals can be extracted simultaneously from their convolutive mixtures. The simulation results have demonstrated the effectiveness of the proposed EVA. However, from the simulation results, one can see that the EVA has such a drawback that it is sensitive to Gaussian noise. Therefore, as a further work, we will propose an EVA having such a property that the BD can be achieved as little insensitive to Gaussian noise as possible.

Acknowledgment

This work is supported by the Research Projects, Grant-IN AIDs, No. 18500146[1] and No. 1850054[2] for Scientific Research of the JSPS.

References

[1] A. Adib, et al., "Source separation contrasts using a reference signal," *IEEE Signal Processing Letters*, Vol. 11, No. 3, pp. 312-315, 2004.

[2] Y. Inouye, "Autoregressive model fitting for multichannel time series of degenerate rank: Limit properties," IEEE Trans. Circuits and Systems, Vol. 32, No. 3, pp. 252-259, 1985.

[3] Y. Inouye and K. Tanebe, "Super-exponential algorithms for multichannel blind deconvolution," *IEEE Trans. Sig. Proc.*, vol. 48, no. 3, pp. 881-888, Mar. 2000.

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

[5] M. Kawamoto, et al., "Eigenvector algorithms using reference signals," Proc. ICASSP 2006, vol. V, pp. 841-844, May 2006.

[6] K. Kohno, et al., "Adaptive super-exponential algorithms for blind deconvolution of MIMO systems," Proc. ISCAS 2004, vol. V, pp. 680-683, May 2004.

[7] K. Kohno, et al., "Robust super-exponential methods for blind equalization of MIMO-IIR systems," Proc. ICASSP 2006, Vol. V, pp. 661-664, 2006.

[8] L. Parra and P. Sajda, "Blind source separation via generalized eigenvalue decomposition," *Journal of Machine Learning*, No. 4, pp, 1261-1269, 2003.

[9] S. Rhioui, et al. "Quadratic MIMO contrast functions for blind source separation in a convolutive contest," Proc. ICA 2006, pp. 230-237, 2006.