Eigenvector Algorithms Incorporated With Reference Systems for Solving Blind Deconvolution of MIMO-IIR Linear Systems

Mitsuru Kawamoto, *Member, IEEE*, Kiyotaka Kohno, *Student Member, IEEE*, and Yujiro Inouye, *Member, IEEE*

***Abstract*—This letter presents an eigenvector algorithm (EVA) for blind deconvolution (BD) of multiple-input multiple-output infinite impulse response (MIMO-IIR) channels (convolutive mixtures), using the idea of reference signals. Differently from the conventional researches on EVAs, the proposed EVA utilizes only one reference signal for recovering all the source signals simultaneously. Computer simulations are presented for demonstrating the effectiveness of the proposed algorithm.**

***Index Terms*—Blind deconvolution, blind signal processing, eigenvector algorithms, multiple-input multiple-output infinite impulse response (MIMO-IIR) channels, reference systems.**

I. Introduction

In this letter, 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) [6], [13]. The first proposal of the EVA was done by Jelonnek *et al.* [6]. 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 [4], [9], and [13], 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 that can solve the BSS problems in the case of MIMO instantaneous mixtures [7].

The EVA in [7] was 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 of our knowledge, have been made by Jelonnek *et al.* (e.g., [6]), Adib *et al.* (e.g., [1]), Rhioui [14], and Castella [3]. 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.* [14] and Castella *et al.* [3] have proposed quadratic MIMO contrast functions for the BSS with convolutive mixtures, and they 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 that can be extracted were used, and in [3], a deflation approach was used, in which for each deflation, a different reference signal was used.

In this letter, it will be shown how the EVA derived by using reference signals works for the BD of the MIMO-IIR channels (see Theorem 1). From the analysis, one can see that the proposed EVA works such that all source signals can be recovered using only one reference signal. This is a novel major result in this letter. Simulation results are presented to show the effectiveness of the proposed EVA.

This letter 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 dag denote, respectively, the transpose, the complex conjugate, the complex conjugate transpose (Hermitian), and the pseudo-inverse operation 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 a MIMO channel with n inputs and m outputs as described by

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

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 $\boldsymbol{v}(t) = \sum_{k=-\infty}^{\infty} \boldsymbol{W}^{(k)}\boldsymbol{y}(t-k) = \sum_{k=-\infty}^{\infty} \boldsymbol{G}^{(k)}\boldsymbol{s}(t-k) +$

Manuscript received February 19, 2007; revised July 9, 2007. This work was supported by the Grant-in-Aids for the Scientific Research by the Ministry of Education, Culture, Sports, Science and Technology of Japan under Grant 18500146 and Grant 18500054. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Arie Yeredor.

M. Kawamoto is with the National Institute of Advanced Industrial Science and Technology (AIST), Ibaraki 305-8568, Japan, and is also with the Laboratory for Biologically Integrative Sensors, Bio-Mimetic Control Research Center, RIKEN, Nagoya 463-0003, Japan (e-mail: m.kawamoto@aist.go.jp).

K. Kohno and Y. Inouye are with the Department of Electronic and Control Systems Engineering, Shimane University, Shimane 690-8504, Japan.

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/LSP.2007.906225

$\sum_{k=-\infty}^{\infty} \boldsymbol{W}^{(k)}\boldsymbol{n}(t-k)$, 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 [8, Fig. 1]. The assumptions on the channel, the source signal, the deconvolver, and the noise are as follows.

1) The transfer function $\boldsymbol{H}(z)$ is stable and has full column rank on the unit circle $|z| = 1$, where the assumption 1) 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.
2) 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 $\gamma_i = \mathrm{cum}\{s_i(t), s_i(t), s_i^*(t), s_i^*(t)\} \neq 0$.
3) The deconvolver $\boldsymbol{W}(z)$ is an FIR channel of sufficient length L so that the truncation effect can be ignored. We define $\boldsymbol{W}(z) := \sum_{k=L_1}^{L_2} \boldsymbol{W}^{(k)}z^k$.
4) 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 $\sigma_{n_j}^2, j = \overline{1,m}$.
5) The two vector sequences $\{\boldsymbol{n}(t)\}$ and $\{\boldsymbol{s}(t)\}$ are mutually statistically independent.

Under 3), the impulse response $\{\boldsymbol{G}^{(k)}\}$ of the cascade system can be written in a vector form as $\tilde{\boldsymbol{g}}_i = \tilde{\boldsymbol{H}}\tilde{\boldsymbol{w}}_i, i = \overline{1,n}$, where $\tilde{\boldsymbol{g}}_i$ is the column vector consisting of the ith output impulse response of the cascade system defined by $\tilde{\boldsymbol{g}}_i := [\boldsymbol{g}_{i1}^T, \boldsymbol{g}_{i2}^T, \ldots, \boldsymbol{g}_{in}^T]^T, \boldsymbol{g}_{ij} := [\cdots, g_{ij}(-1), g_{ij}(0), g_{ij}(1), \ldots]^T, j = \overline{1,n}$, 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 as in [8], 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 $[\boldsymbol{H}_{ij}]_{lr} := h_{ji}(l-r), l = 0, \pm1, \pm2, \ldots, r = \overline{L_1, L_2}$, 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

$$[\tilde{\boldsymbol{g}}_1, \ldots, \tilde{\boldsymbol{g}}_n] = \tilde{\boldsymbol{H}}[\tilde{\boldsymbol{w}}_1, \ldots, \tilde{\boldsymbol{w}}_n] = \left[\tilde{\boldsymbol{\delta}}_1, \ldots, \tilde{\boldsymbol{\delta}}_n\right]\boldsymbol{P} \tag{2}$$

where $\boldsymbol{P}$ is an $n \times n$ permutation matrix, and $\tilde{\boldsymbol{\delta}}_i$ is the n-block column vector defined by $\tilde{\boldsymbol{\delta}}_i := [\boldsymbol{\delta}_{i1}^T, \boldsymbol{\delta}_{i2}^T, \ldots, \boldsymbol{\delta}_{in}^T]^T, i = \overline{1,n}, \boldsymbol{\delta}_{ij} := \hat{\boldsymbol{\delta}}_i$, for $i = j$, otherwise $(\cdots, 0, 0, 0, \ldots)^T$.

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)

In this section, we assume that there is no noise $\boldsymbol{n}(t)$ in the output $\boldsymbol{y}(t)$, and then, we analyze eigenvector algorithms for the MIMO channel (1). We should note that the noise $\boldsymbol{n}(t)$ in the output $\boldsymbol{y}(t)$ is taken into account in computer simulations (see Section IV).

A. Analysis of Eigenvector Algorithms With Reference Signals for MIMO-IIR Channels

In order to solve the BD problem, the following cross-cumulant between $v_i(t)$ and a reference signal $x(t)$, which is also utilized in [3], is defined as follows:

$$D_{v_i x} = \mathrm{cum}\{v_i(t), v_i^*(t), x(t), x^*(t)\} \tag{3}$$

where $v_i(t)$ is the ith element of $\boldsymbol{v}(t)$ that is the output of the deconvolver, and the reference signal $x(t)$ is given by $\boldsymbol{f}^T(z)\boldsymbol{y}(t)$, using an appropriate filter $\boldsymbol{f}(z)$ (see [8, Fig. 1]). The filter $\boldsymbol{f}(z)$ is called a *reference system*. Note that the input of the reference system is $\boldsymbol{y}(t)$. Let $\boldsymbol{a}(z) := \boldsymbol{H}^T(z)\boldsymbol{f}(z) = [a_1(z), a_2(z), \ldots, a_n(z)]^T$, and 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.* [6] have shown in the single-input case that by the Lagrangian method, the maximization of $|D_{v_i x}|$ under $\sigma_{v_i}^2 = \sigma_{s_{\rho_i}}^2$ leads to a closed-form expressed as a generalized eigenvector problem, 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 $\{1, 2, \ldots, n\}$ such that the set $\{\rho_1, \rho_2 \ldots, \rho_n\}$ is a permutation of the set $\{1, 2, \ldots, n\}$. In our case, $D_{v_i x}$ and $\sigma_{v_i}^2$ can be expressed in terms of the vector $\tilde{\boldsymbol{w}}_i$ as, respectively, $D_{v_i x} = \tilde{\boldsymbol{w}}_i^H\tilde{\boldsymbol{B}}\tilde{\boldsymbol{w}}_i$ and $\sigma_{v_i}^2 = \tilde{\boldsymbol{w}}_i^H\tilde{\boldsymbol{R}}\tilde{\boldsymbol{w}}_i$, 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 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

$$\tilde{\boldsymbol{y}}(t) := \left[\boldsymbol{y}_1^T(t), \boldsymbol{y}_2^T(t), \ldots, \boldsymbol{y}_m^T(t)\right]^T \in \boldsymbol{C}^{mL} \tag{4}$$

where $\boldsymbol{y}_j(t) := [y_j(t-L_1), y_j(t-L_1-1), \ldots, y_j(t-L_2)]^T \in \boldsymbol{C}^L, j = \overline{1,m}$. Therefore, by the similar way as in [6], the maximization of $|D_{v_i x}|$ under $\sigma_{v_i}^2 = \sigma_{s_{\rho_i}}^2$ leads to the following generalized eigenvector problem:

$$\tilde{\boldsymbol{B}}\tilde{\boldsymbol{w}}_i = \lambda_i\tilde{\boldsymbol{R}}\tilde{\boldsymbol{w}}_i. \tag{5}$$

Moreover, Jelonnek *et al.* have shown in [6] 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, which is referred to as an 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 (5) are different from those proposed in [6]. In this letter, under the assumption that any reference system $\boldsymbol{f}(z)$ is used, we want to show how the eigenvector algorithm (5) works for the BD of the MIMO-IIR channel (1). To this end, we use the following equalities:

$$\tilde{\boldsymbol{R}} = \tilde{\boldsymbol{H}}^H\tilde{\boldsymbol{\Sigma}}\tilde{\boldsymbol{H}}, \tilde{\boldsymbol{B}} = \tilde{\boldsymbol{H}}^H\tilde{\boldsymbol{\Lambda}}\tilde{\boldsymbol{H}} \tag{6}$$

where $\tilde{\boldsymbol{\Sigma}}$ is the block diagonal matrix defined by $\tilde{\boldsymbol{\Sigma}} := \text{block-diag}\{\boldsymbol{\Sigma}_1, \boldsymbol{\Sigma}_2, \ldots, \boldsymbol{\Sigma}_n\}, \boldsymbol{\Sigma}_i := \mathrm{diag}\{\cdots, \sigma_{s_i}^2, \sigma_{s_i}^2, \sigma_{s_i}^2, \ldots\}, i = \overline{1,n}$, and $\tilde{\boldsymbol{\Lambda}}$ is the block diagonal matrix defined by $\tilde{\boldsymbol{\Lambda}} := \text{block-diag}\{\boldsymbol{\Lambda}_1, \boldsymbol{\Lambda}_2, \ldots, \boldsymbol{\Lambda}_n\}, \boldsymbol{\Lambda}_i := \mathrm{diag}\{\cdots, |a_i(-1)|^2\gamma_i, |a_i(0)|^2\gamma_i, |a_i(1)|^2\gamma_i, \ldots\}, i = \overline{1,n}$. Since both $\tilde{\boldsymbol{\Sigma}}$ and $\tilde{\boldsymbol{\Lambda}}$ become diagonal, (6) shows that the two matrices $\tilde{\boldsymbol{R}}$ and $\tilde{\boldsymbol{B}}$ are simultaneously diagonalizable. Here, let the eigenvalues of the diagonal matrix $\tilde{\boldsymbol{\Sigma}}^{-1}\tilde{\boldsymbol{\Lambda}}$ be denoted as $\lambda_i(k) := |a_i(k)|^2\gamma_i/\sigma_{s_i}^2, i = \overline{1,n}, k \in Z$. We put the following assumption on the eigenvalues $\lambda_i(k)'s$: 6) 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 1) through 6), 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 (2).

Outline of the Proof: Based on (5), we consider the following eigenvector problem:

$$\tilde{\boldsymbol{R}}^\dagger\tilde{\boldsymbol{B}}\tilde{\boldsymbol{w}}_i = \lambda_i\tilde{\boldsymbol{w}}_i. \tag{7}$$

Then, from (6), (7) becomes

$$(\tilde{\boldsymbol{H}}^H\tilde{\boldsymbol{\Sigma}}\tilde{\boldsymbol{H}})^\dagger\tilde{\boldsymbol{H}}^H\tilde{\boldsymbol{\Lambda}}\tilde{\boldsymbol{H}}\tilde{\boldsymbol{w}}_i = \lambda_i\tilde{\boldsymbol{w}}_i. \tag{8}$$

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

$$(\tilde{\boldsymbol{H}}^H\tilde{\boldsymbol{\Sigma}}\tilde{\boldsymbol{H}})^\dagger = \tilde{\boldsymbol{H}}^\dagger\tilde{\boldsymbol{\Sigma}}^\dagger\tilde{\boldsymbol{H}}^{H\dagger}, \tilde{\boldsymbol{H}}^{H\dagger}\tilde{\boldsymbol{H}}^H = \boldsymbol{I} \tag{9}$$

which are shown in [12] along with their proofs. Then, it follows the following form from (8):

$$\tilde{\boldsymbol{H}}^\dagger\tilde{\boldsymbol{\Sigma}}^{-1}\tilde{\boldsymbol{\Lambda}}\tilde{\boldsymbol{H}}\tilde{\boldsymbol{w}}_i = \lambda_i\tilde{\boldsymbol{w}}_i. \tag{10}$$

Multiplying (10) by $\tilde{\boldsymbol{H}}$ from the left side and using (9), (10) becomes

$$\tilde{\boldsymbol{\Sigma}}^{-1}\tilde{\boldsymbol{\Lambda}}\tilde{\boldsymbol{H}}\tilde{\boldsymbol{w}}_i = \lambda_i\tilde{\boldsymbol{H}}\tilde{\boldsymbol{w}}_i. \tag{11}$$

$\tilde{\boldsymbol{\Sigma}}^{-1}\tilde{\boldsymbol{\Lambda}}$ is a diagonal matrix with diagonal elements $\lambda_i(k), i = \overline{1,n}$ and $k \in Z$, and thus, (7) and (11) 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: $\lim_{L\to\infty}(\mathrm{rank}\tilde{\mathbf{R}})/L = n$, which is shown in [10] and whose proof is found in [5]. Using this fact, the other remaining eigenvalues of $\tilde{\boldsymbol{R}}^\dagger\tilde{\boldsymbol{B}}$ are all equal to zero. From the assumption 6), the n nonzero eigenvalues $\lambda_i(k) \neq 0, i = \overline{1,n}$, obtained by (11), 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 (7) become n solutions of the vectors $\tilde{\boldsymbol{w}}_i$ satisfying (2).

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 6) that there exist nL nonzero eigenvalues of $\tilde{\boldsymbol{R}}^\dagger\tilde{\boldsymbol{B}}$ that 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 $\tilde{\boldsymbol{R}}^\dagger\tilde{\boldsymbol{B}}$ that are approximately equal to zero.

Remark 2: The proposed EVA is closely related to the joint diagonalization of square matrices (e.g., [2]).

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 (2), 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:

$$\tilde{\boldsymbol{B}}\tilde{\boldsymbol{R}}^\dagger\hat{\boldsymbol{w}}_i = \hat{\lambda}_i\hat{\boldsymbol{w}}_i \tag{12}$$

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 (12) 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 (12) correspond (or are equal) to the row vectors of $\tilde{\boldsymbol{H}}$ up to constants. The proof of the mentioned above is given below: Substituting (6) into (12), we obtain

$$\tilde{\boldsymbol{H}}^H\tilde{\boldsymbol{\Lambda}}\tilde{\boldsymbol{H}}(\tilde{\boldsymbol{H}}^H\tilde{\boldsymbol{\Sigma}}\tilde{\boldsymbol{H}})^\dagger\hat{\boldsymbol{w}}_i = \hat{\lambda}_i\hat{\boldsymbol{w}}_i. \tag{13}$$

By the similar way to (10), (13) becomes

$$\tilde{\boldsymbol{H}}^H\tilde{\boldsymbol{\Lambda}}\tilde{\boldsymbol{\Sigma}}^{-1}\tilde{\boldsymbol{H}}^{H\dagger}\hat{\boldsymbol{w}}_i = \hat{\lambda}_i\hat{\boldsymbol{w}}_i. \tag{14}$$

Multiplying (14) by $\tilde{\boldsymbol{H}}^{H\dagger}$ from the left side,(14) becomes

$$\tilde{\boldsymbol{\Lambda}}\tilde{\boldsymbol{\Sigma}}^{-1}\tilde{\boldsymbol{H}}^{H\dagger}\hat{\boldsymbol{w}}_i = \hat{\lambda}_i\tilde{\boldsymbol{H}}^{H\dagger}\hat{\boldsymbol{w}}_i. \tag{15}$$

Let $\hat{\boldsymbol{g}}_i := \tilde{\boldsymbol{H}}^{H\dagger}\hat{\boldsymbol{w}}_i$, and then, (15) 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

$$\tilde{\boldsymbol{H}}^H\hat{\boldsymbol{g}}_i = \tilde{\boldsymbol{H}}^H\tilde{\boldsymbol{H}}^{H\dagger}\hat{\boldsymbol{w}}_i. \tag{16}$$

We obtain from (13) 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 $\tilde{\boldsymbol{H}}^H\tilde{\boldsymbol{H}}^{H\dagger}\tilde{\boldsymbol{H}}^H = \tilde{\boldsymbol{H}}^H$,(16) gives $\tilde{\boldsymbol{H}}^H\hat{\boldsymbol{g}}_i = \tilde{\boldsymbol{H}}^H\tilde{\boldsymbol{H}}^{H\dagger}\hat{\boldsymbol{w}}_i = \tilde{\boldsymbol{H}}^H\tilde{\boldsymbol{H}}^{H\dagger}\tilde{\boldsymbol{H}}^H\hat{\boldsymbol{\xi}}_i = \tilde{\boldsymbol{H}}^H\hat{\boldsymbol{\xi}}_i = \hat{\boldsymbol{w}}_i$, 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 above (2)] 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 of the channel, which is denoted by K, 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 (7) corresponding to nL nonzero eigenvalues into n sets of L eigenvectors whose pairs have almost the same correlations for all pairs of eigenvectors of each set. There remain $(m-n)L$ eigenvectors corresponding to the remaining $(m-n)L$ eigenvalues that are approximately zero. Thus, we propose a tentative procedure of finding n eigenvectors satisfying (2) as follows.

1) Set $N_{\mathrm{ite}} = 1$ (where N_{ite} denotes the number of iterations from the beginning less than $n+1$).
2) Calculate the eigenvectors $\tilde{\boldsymbol{w}}_i, i = \overline{1,mL}$ of $\tilde{\boldsymbol{R}}^\dagger\tilde{\boldsymbol{B}}$ and the eigenvectors $\hat{\boldsymbol{w}}_i, i = \overline{1,mL}$ of $\tilde{\boldsymbol{B}}\tilde{\boldsymbol{R}}^\dagger$. Then, select $\tilde{\boldsymbol{w}}_{i_0}$ and $\hat{\boldsymbol{w}}_{i_0}$ corresponding to the maximum magnitude eigenvalue $|\hat{\lambda}_{i_0}|$ among $|\hat{\lambda}_i|$'s for $i = \overline{1,(m+1-N_{\mathrm{ite}})L}$ (where

we should note $\tilde{\lambda}_i = \hat{\lambda}_i$ for $i = \overline{1, (m+1-N_{\text{ite}})L}$ in theory).

3) Calculate the magnitudes of the correlations of all the pairs of $\hat{\boldsymbol{w}}_{i_0}$ and $\hat{\boldsymbol{w}}_i$ for $i = \overline{1, (m+1-N_{\text{ite}})L}$. Then, select the Lth largest magnitude correlation of all the $(m+1-N_{\text{ite}})L$ magnitude correlations calculated in 2.
4) Remove L eigenvalues $\hat{\lambda}_i$'s corresponding to the L magnitude correlations larger than or equal to the Lth largest magnitude one selected in 3). Then save the $(m-N_{\text{ite}})L$ remaining eigenvalues $\hat{\lambda}_i$'s for finding other eigenvectors satisfying (2).
5) Put $N_{\text{ite}} = N_{\text{ite}} + 1$ and stock the $\tilde{\boldsymbol{w}}$ obtained in 2). If $N_{\text{ite}} = n+1$, stop the iterations; otherwise, go to 2).

Therefore, the n eigenvectors $\tilde{\boldsymbol{w}}_{i_0}$'s stocked in step 5) are the n solutions in (2)

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 (see [8]). The parameters L_1 and L_2 in $\boldsymbol{W}(z)$ were set to be 0 and 11, respectively. As a measure of performances, we used the *multichannel inter-symbol interference* (M_{ISI}) [11]. For comparison, the algorithm proposed by Castella *et al.* (CRMPA) was used [3]. Fig. 1 shows the performance results obtained by our EVA (black lines) and the CRMPA (gray lines), 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. 1 was the average of the performances obtained by 50 independent Monte Carlo runs. In each Monte Carlo run, the eigenvectors calculated from each algorithm were obtained by ten updates, where in each update, the estimates of $\tilde{\boldsymbol{R}}$ and $\tilde{\boldsymbol{B}}$ were modified by 5000 data samples. For the two solid black and gray lines in Fig. 1, $s_1(t)$ and $s_2(t)$ were 4-QAM and 8-QAM, respectively, and the reference signals of our EVA and the CRMPA were, respectively, $f_2(2)y_2(t-2)$ and $v_i(t)$ for $i = \overline{1,2}$, where the parameter $f_2(2)$ was randomly chosen from a Gaussian distribution with zero mean and unit variance. It can be seen from Fig. 1 that both performances are almost the same. However, if the reference signal and the source signal were changed for our EVA, the performances corresponding to such cases were changed (see the black dotted and dashed lines), where in black dotted and dashed lines, the reference signals were set to be $\sum_{i=1}^{3} f_i(2)y_i(t-2)$ and $f_2(2)y_2(t-2)$, respectively, and two source signals were 2-PAM that takes one of two values, -1 and 1 with equal probability $1/2$. On the other hand, the CRMPA hardly changes the performance (see gray dashed line in Fig. 1), even for the case that the source signals were 2-PAM. For the interested reader, the details on the influence of the reference signals can be found in [3]. Moreover, since a deflation method was adopted in the CRMPA [3], whereas our EVA can recover all source signals simultaneously without using a deflation method, there is a slight difference for their computational times [For example, under the three conditions that the source signals are the QAM signals, that the ability of the machine is Dual Core Processor (2.59 GHz, 2.0 GB RAM), and that the function (tic, toc) of MATLAB ver.7.3 is used, the computational times measured for one update are, respectively, 2.22 [s] (for our EVA) and 3.53 [s] (for the CRMPA)]. From all the results, we conclude that our EVA has almost the same ability as the CRMPA, although the computational time of our EVA is a little shorter than the CRMPA. Moreover, our EVA has a possibility of obtaining better performances than the CRMPA, choosing an appropriate reference signal.

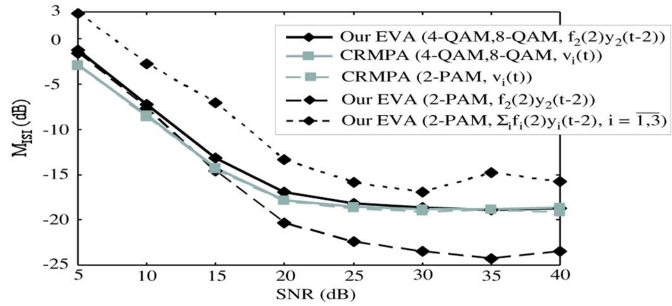

Fig. 1. Performances of our EVA and the CRMPA with varying SNR levels, in the cases of 5000 data samples.

V. Conclusion

We have proposed an EVA for solving the BD problem. 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. In a forthcoming paper, we will propose an EVA having such a property that the BD can be achieved with as little insensitivity to Gaussian noise as possible.

References

[1] A. Adib *et al.*, “Source separation contrasts using a reference signal,” *IEEE Signal Process. Lett.*, vol. 11, no. 3, pp. 312–315, Mar. 2004.
[2] J.-F. Cardoso and A. Souloumiac, “Jacobi angles for simultaneous diagonalization,” *SIAM J. Matrix Anal. Appl.*, vol. 17, pp. 161–164, 1996.
[3] M. Castella *et al.*, “Quadrativ higher-order criteria for iterative blind separation of a MIMO convolutive mixture of sources,” *IEEE Trans. Signal Process.*, vol. 55, no. 1, pp. 218–232, Jan. 2007.
[4] R. M. Clemente *et al.*, “Eigendecomposition of self-tuned cumulant-matrices for blind source separation,” *Signal Process.*, vol. 84, no. 7, pp. 1201–1211, Jul. 2004.
[5] Y. Inouye, “Autoregressive model fitting for multichannel time series of degenerate rank: Limit properties,” *IEEE Trans. Circuits Syst.*, vol. 32, no. 3, pp. 252–259, Mar. 1985.
[6] B. Jelonnek and K. D. Kammeyer, “A closed-form solution to blind equalization,” *Signal Process.*, vol. 36, no. 3, pp. 251–259, 1994.
[7] M. Kawamoto *et al.*, “Eigenvector algorithms using reference signals,” in *Proc. ICASSP 2006*, May 2006, vol. V, pp. 841–844.
[8] M. Kawamoto *et al.*, “Eigenvector algorithms for blind deconvolution of MIMO-IIR systems,” *Proc. ISCAS 2007*, pp. 3486–3489, May 2007.
[9] J.-Y. Ko *et al.*, “Eigenvalue problem approach to the blind source separation: Optimization with a reference signal,” *Phys. Rev. E*, vol. 58, no. 4, pp. 4872–4882, May 2004.
[10] K. Kohno *et al.*, “Adaptive super-exponential algorithms for blind deconvolution of MIMO systems,” in *Proc. ISCAS 2004*, May 2004, vol. V, pp. 680–683.
[11] K. Kohno *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.
[12] K. Kohno *et al.*, “Robust super-exponential methods for blind equalization of MIMO-IIR systems,” in *Proc. ICASSP 2006*, 2006, vol. V, pp. 661–664.
[13] L. Parra and P. Sajda, “Blind source separation via generalized eigenvalue decomposition,” *J. Mach. Learn.*, no. 4, pp. 1261–1269, 2003.
[14] S. Rhioui *et al.*, “Quadratic MIMO contrast functions for blind source separation in a convolutive contest,” in *Proc. ICA 2006*, 2006, pp. 230–237.