

EIGENVECTOR ALGORITHMS USING REFERENCE SIGNALS

Mitsuru Kawamoto^{1†}, Kiyotaka Kohno², and Yujiro Inouye²

¹AIST, Japan, ²Shimane University, Japan

ABSTRACT

This paper presents an eigenvector algorithm (EVA) derived from a criterion using reference signals, in which the EVA is applied to the blind source separation (BSS) of instantaneous mixtures. The proposed EVA works such that source signals are simultaneously separated from their mixtures. This is a new result, which has not been clarified by the conventional researches. Simulation results show the validity of the proposed EVA.

1. INTRODUCTION

This paper deals with the blind source separation (BSS) problem for a multiple-input and multiple-output (MIMO) static system driven by independent source signals. To solve this problem, reference signals are used. Researches on the BSS problem by using the idea of reference signals, to our best knowledge, have been made by Jelonnek et al. [4, 5] and Adib et al. [1]. Jelonnek et al. have proposed an eigenvector algorithm (EVA) derived from a criterion using reference signals, in order to solve blind equalization of single-input and single-output (SISO) systems. Adib et al. have proposed contrast functions for solving the BSS problem, in which reference signals are included into the contrast functions, but they have not derived explicit algorithms for solving the BSS problem from the contrast functions.

In this paper, the EVA derived from a criterion with reference signals is used for solving the BSS problem of MIMO static systems, and then it will be shown that the EVA works such that source signals are simultaneously separated from their mixtures. Simulation results show that the proposed EVA works successfully to solve the BSS problem and thanks to its property, provides better performances, compared with the super-exponential algorithm using deflation methods.

2. PROBLEM FORMULATION

Throughout this paper, let us consider the following MIMO static system with n inputs and m outputs:

$$\mathbf{y}(t) = \mathbf{H}\mathbf{s}(t) + \mathbf{n}(t), \quad (1)$$

[†]The author is also with Bio-Mimetic Control Research Center, RIKEN, Nagoya, Japan.

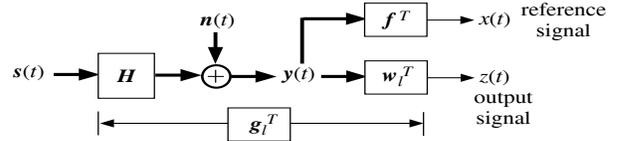


Fig. 1. The composite system of an unknown system and a filter, and reference system.

where $\mathbf{y}(t)$ represents an m -column output vector called the *observed signal*, $\mathbf{s}(t)$ represents an n -column input vector called the *source signal*, \mathbf{H} is an $m \times n$ matrix, $\mathbf{n}(t)$ represents an m -column noise vector. It can be regarded as a linear mixture model with additive noise.

To achieve the blind source separation (BSS) for the system (1), the following n filters, which are m -input single-output (MISO) static systems driven by the observed signals, are used:

$$z_l(t) = \mathbf{w}_l^T \mathbf{y}(t), \quad l = 1, 2, \dots, n, \quad (2)$$

where $z_l(t)$ is the l th output of the filter, $\mathbf{w}_l = [w_{l1}, w_{l2}, \dots, w_{lm}]^T$ is an m -column vector representing the m coefficients of the filter. Substituting (1) into (2), we obtain

$$\begin{aligned} z_l(t) &= \mathbf{w}_l^T \mathbf{H} \mathbf{s}(t) + \mathbf{w}_l^T \mathbf{n}(t) \\ &= \mathbf{g}_l^T \mathbf{s}(t) + \mathbf{w}_l^T \mathbf{n}(t), \quad l = 1, 2, \dots, n, \end{aligned} \quad (3)$$

where $\mathbf{g}_l = [g_{l1}, g_{l2}, \dots, g_{ln}]^T := \mathbf{H}^T \mathbf{w}_l$ is an n -column vector. The BSS problem considered in this paper can be formulated as follows: Find n filters \mathbf{w}_l 's denoted by $\tilde{\mathbf{w}}_l$'s satisfying the following condition, without the knowledge of \mathbf{H} .

$$\tilde{\mathbf{g}}_l = \mathbf{H}^T \tilde{\mathbf{w}}_l = \tilde{\boldsymbol{\delta}}_l, \quad l = 1, 2, \dots, n, \quad (4)$$

where $\tilde{\boldsymbol{\delta}}_l$ is an n -column vector whose elements $\tilde{\delta}_{lr}$ ($r = 1, 2, \dots, n$) are equal to zero except for ρ_l th element, that is, $\tilde{\delta}_{lr} = c_l \delta(r - \rho_l)$, $r = 1, 2, \dots, n$. Here, $\delta(t)$ is the Kronecker delta function, c_l is a number standing for a scale change, and ρ_l 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 $\{1, 2, \dots, n\}$.

To solve the BSS problem, we put the following assumptions on the system and the source signals.

A1) The matrix \mathbf{H} in (1) is an $m \times n$ ($m \geq n$) matrix and has full column rank.

A2) The input sequence $\{\mathbf{s}(t)\}$ is a zero-mean, non-Gaussian vector stationary process whose element processes $\{s_i(t)\}$,

$i = 1, 2, \dots, n$, are mutually statistically independent and have nonzero variance, $\sigma_{s_i}^2 = E[s_i^2(t)] \neq 0$ and nonzero fourth-order cumulants, γ_i defined as

$$\gamma_i = \text{cum}\{s_i(t), s_i(t), s_i(t), s_i(t)\} \neq 0 \text{ for } i = 1, 2, \dots, n, \quad (5)$$

A3) The noise signal sequence $\{\mathbf{n}(t)\}$ is a zero-mean, Gaussian vector stationary process whose element processes $\{n_i(t)\}$, $i = 1, 2, \dots, m$, are mutually statistically independent.

A4) The two vector sequences $\{\mathbf{n}(t)\}$ and $\{\mathbf{s}(t)\}$ are mutually statistically independent.

It is assumed for the sake of simplicity in this paper that all the signals and all the systems are real-valued.

3. EIGENVECTOR ALGORITHM (EVA) WITH REFERENCE SIGNALS FOR MIMO STATIC SYSTEMS

In this section, we assume that there is no noise $\mathbf{n}(t)$ in the output $\mathbf{y}(t)$, and then analyze eigenvector algorithms for MIMO static systems. Under this assumption, to solve the BSS problem, the following cross-cumulant between $z_l(t)$ and a reference signal $x(t)$ (see Fig. 1) is defined:

$$C_{zx} = \text{cum}\{z_l(t), z_l(t), x(t), x(t)\} \quad (6)$$

where the reference signal $x(t)$ is given by $\mathbf{f}^T \mathbf{y}(t) = \mathbf{f}^T \mathbf{H} \mathbf{s}(t) = \mathbf{a}^T \mathbf{s}(t)$ ($\mathbf{a}^T := \mathbf{f}^T \mathbf{H}$ is a vector whose elements are a_1, a_2, \dots, a_n), using an appropriate filter \mathbf{f} . The filter \mathbf{f} is called a *reference system*. Moreover we define the constraint $\sigma_{z_l}^2 = \sigma_{s_{\rho_l}}^2$, where $\sigma_{z_l}^2$ and $\sigma_{s_{\rho_l}}^2$ denote the variances of the output $z_l(t)$ and a source signal $s_{\rho_l}(t)$, respectively. Adib, et al. [1] have shown that the BSS can be achieved by maximizing $|C_{zx}|$ in (6) under the constraint, but they have not proposed any algorithm for achieving this idea. In the single-input case, Jelonnek et al. [4, 5] have shown that by the Lagrangian method, the maximization of $|C_{zx}|$ under $\sigma_{z_l}^2 = \sigma_{s_{\rho_l}}^2$ leads to a closed-form expression as the following generalized eigenvector problem:

$$\mathbf{C}_{yx} \mathbf{w}_l = \lambda \mathbf{R} \mathbf{w}_l \quad (7)$$

Then they utilize the facts that C_{zx} and $\sigma_{z_l}^2$ can be expressed in terms of the vector \mathbf{w}_l as, respectively,

$$C_{zx} = \mathbf{w}_l^T \mathbf{C}_{yx} \mathbf{w}_l, \quad (8)$$

$$\sigma_{z_l}^2 = \mathbf{w}_l^T \mathbf{R} \mathbf{w}_l, \quad (9)$$

where \mathbf{C}_{yx} is a matrix whose (i, j) element is calculated by $\text{cum}\{y_i(t), y_j(t), x(t), x(t)\}$ and $\mathbf{R} = E[\mathbf{y}(t) \mathbf{y}^T(t)]$ is the covariance matrix of m -column vector $\mathbf{y}(t)$. Moreover, they have shown that the eigenvector corresponding to the maximum eigenvalue of $\mathbf{R}^\dagger \mathbf{C}_{yx}$ becomes the solution of the blind equalization problem in [4, 5], which is referred to as an *eigenvector algorithm* (EVA). However, the algorithm proposed by

Jelonnek et al. is for SISO or SIMO infinite impulse response channel. Therefore, we want to show how the eigenvector algorithm (7) works for the BSS in the case of the MIMO static system. To this end, we use the following equalities:

$$\mathbf{R} = \mathbf{H} \mathbf{\Sigma} \mathbf{H}^T, \quad (10)$$

$$\mathbf{C}_{yx} = \mathbf{H} \mathbf{\Lambda} \mathbf{H}^T, \quad (11)$$

where $\mathbf{\Sigma}$ is a diagonal matrix whose elements are $\sigma_{s_i}^2$, $i = 1, 2, \dots, n$ and $\mathbf{\Lambda}$ is a diagonal matrix whose elements are $a_i^2 \gamma_i$ ($i = 1, 2, \dots, n$). Then we obtain the following theorem.

Theorem 1 *Suppose the values $a_i^2 \gamma_i / \sigma_{s_i}^2$, $i = 1, 2, \dots, n$ are all nonzero and distinct. If the noise $\mathbf{n}(t)$ is absent in (1), then the n eigenvectors corresponding to n nonzero eigenvalues of $\mathbf{R}^\dagger \mathbf{C}_{yx}$ become the vectors $\tilde{\mathbf{w}}_l$'s satisfying (4), where the symbol \dagger denotes the pseudo-inverse operation of a matrix.*

Proof: Based on (7), we consider the following eigenvector problem:

$$\mathbf{R}^\dagger \mathbf{C}_{yx} \mathbf{w}_l = \lambda \mathbf{w}_l. \quad (12)$$

Then, from (10) and (11), (12) becomes

$$\mathbf{H}^T \mathbf{\Sigma}^{-1} \mathbf{H}^\dagger \mathbf{H} \mathbf{\Lambda} \mathbf{H}^T \mathbf{w}_l = \lambda \mathbf{w}_l. \quad (13)$$

Since \mathbf{H} has full column rank, using a property of the pseudo-inverse operation ([8], p. 433), we obtain

$$\mathbf{H}^T \mathbf{\Sigma}^{-1} \mathbf{\Lambda} \mathbf{H}^T \mathbf{w}_l = \lambda \mathbf{w}_l. \quad (14)$$

Multiplying (14) by \mathbf{H}^T from left side and using a property of the pseudo-inverse operation again, (14) becomes

$$\mathbf{\Sigma}^{-1} \mathbf{\Lambda} \mathbf{H}^T \mathbf{w}_l = \lambda \mathbf{H}^T \mathbf{w}_l. \quad (15)$$

By noting that $\mathbf{\Sigma}^{-1} \mathbf{\Lambda}$ is a diagonal matrix whose elements, $a_i^2 \gamma_i / \sigma_{s_i}^2$, $i = 1, 2, \dots, n$, are all nonzero and distinct, if $\mathbf{g}_l := \mathbf{H}^T \mathbf{w}_l \neq 0$, then the eigenvectors \mathbf{g}_l obtained from (15) become the vectors $\tilde{\mathbf{g}}_l$ satisfying (4). Namely, the n eigenvectors \mathbf{w}_l corresponding to n nonzero eigenvalues of $\mathbf{R}^\dagger \mathbf{C}_{yx}$ obtained from (12) become the vectors $\tilde{\mathbf{w}}_l$ satisfying (4).

Remark 1 *In order to use Theorem 1, the reference signal $x(t)$ contains nonzero contributions a_i 's from all source signals $s_i(t)$'s. This is the case except for pathological cases. From Theorem 1, it can be seen that by all the n eigenvectors corresponding to n nonzero eigenvalues of $\mathbf{R}^\dagger \mathbf{C}_{yx}$, all source signals can be separated from the output $\mathbf{y}(t)$. This is a novel result which has not been shown in the conventional researches. Moreover, it can be seen from (15) that even if the fourth-order cumulants γ_i have different signs from each other, the vector $\tilde{\mathbf{w}}_l$ satisfying (4) can be obtained. This fact will be confirmed by computer simulations in Section 5.*

Remark 2 *In this section, we assume that there are no noises in the output signals. We are able to show such an eigenvector algorithm that the solutions (4) can be obtained, even if the noise $\mathbf{n}(t)$ is presented in the output $\mathbf{y}(t)$. On the details of this eigenvector algorithm, see [7].*

4. DISCUSSION

In this section, let us compare the proposed EVA with the well-known method obtained by the following constrained maximization problem [9]:

$$\begin{aligned} & \text{Maximize } |C_z| = |\text{cum}\{z_l(t), z_l(t), z_l(t), z_l(t)\}|, \\ & \text{subject to } \sigma_{z_l}^2 = \sigma_{s_{\rho_l}}^2. \end{aligned} \quad (16)$$

Then, by the similar way in (8), C_z can be expressed in terms of the vector \mathbf{w}_l as

$$C_z = \mathbf{w}_l^T \mathbf{C}_{yz} \mathbf{w}_l, \quad (17)$$

where \mathbf{C}_{yz} is an $m \times m$ matrix whose (r_1, r_2) th element is $\text{cum}\{y_{r_1}(t), y_{r_2}(t), z_l(t), z_l(t)\}$, which can be expressed as

$$\mathbf{C}_{yz} = \mathbf{H} \mathbf{\hat{\Lambda}} \mathbf{H}^T, \quad (18)$$

where $\mathbf{\hat{\Lambda}}$ is a diagonal matrix with the diagonal elements $g_{li}^2 \gamma_i$, $i = 1, 2, \dots, n$.

From (17), under the condition that $z_l(t)$'s in \mathbf{C}_{yz} are fixed, by the Lagrangian method, (16) leads to the following generalized eigenvector problem:

$$\mathbf{C}_{yz} \mathbf{w}_l = \lambda \mathbf{R} \mathbf{w}_l. \quad (19)$$

This equation (19) is similar to (7), but by modifying (19) with the similar way to (12) through (15), one can recognize the difference between (7) and (19). By the similar way to (12) through (15), (19) can be modified as

$$\Sigma^{-1} \mathbf{\hat{\Lambda}} \mathbf{H}^T \mathbf{w}_l = \lambda \mathbf{H}^T \mathbf{w}_l. \quad (20)$$

Since the diagonal elements of $\mathbf{\hat{\Lambda}}$ include g_{li} , $i = 1, 2, \dots, n$, which are the elements of the vector \mathbf{g}_l , if the vector $\tilde{\mathbf{w}}_l$ satisfying (4) is obtained by solving (20), the diagonal elements of $\Sigma^{-1} \mathbf{\hat{\Lambda}}$ become zero except for ρ_l th diagonal element, that is, $\Sigma^{-1} \mathbf{\hat{\Lambda}} = \text{diag}\{0, \dots, 0, g_{l\rho_l}^2 \gamma_{\rho_l} / \sigma_{\rho_l} (\rho_l \text{th element}), 0, \dots, 0\}$. This means that if the algorithm (19) is iteratively used to estimate the vector $\tilde{\mathbf{w}}_l$ in (4) with high accuracy, only one source signal can be separated from $\mathbf{y}(t)$. As an iterative algorithm derived from (19), there exists the following two-step iterative algorithm with respect to \mathbf{w}_l :

$$\mathbf{w}_l^{[1]} = \mathbf{R}^\dagger \mathbf{d}_l, \quad (21)$$

$$\mathbf{w}_l = \mathbf{w}_l^{[1]} / \sqrt{\mathbf{w}_l^{[1]T} \mathbf{R} \mathbf{w}_l^{[1]}}, \quad (22)$$

where $\mathbf{d}_l := (1/\lambda) \mathbf{C}_{yz} \mathbf{w}_l$, which is assumed to be calculated by (26). This is nothing less than the *super-exponential algorithm* (SEA) [10]. Indeed, the SEA is used to separate source signals one by one [3], that is, deflation methods are need to separate all source signals from their mixtures. On the other hand, the proposed EVA (7) can work such that all source signals are separated simultaneously from $\mathbf{y}(t)$, even if the iterative calculations are carried out, because the diagonal elements of $\mathbf{\hat{\Lambda}}$ are of fixed values. This is a big difference between the algorithm (21)-(22) and the proposed EVA.

5. 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 \mathbf{H} was set to be a 4×3 matrix, that is, a three-input four-output system:

$$\mathbf{H} = \begin{bmatrix} 1.0 & 0.4 & 0.6 \\ 0.7 & 1.0 & -0.3 \\ 0.2 & -0.5 & 1.0 \\ -0.45 & 0.25 & 0.7 \end{bmatrix}. \quad (23)$$

The three inputs $s_i(t)$ ($i = 1, 2, 3$) of the system \mathbf{H} were two sub-Gaussian signals and one super-Gaussian signal, in which each sub-Gaussian signal takes one of two values, -1 and 1 with equal probability $1/2$ and the super-Gaussian signal takes -2 , 2 , and 0 with probabilities $1/8$, $1/8$, and $6/8$, respectively. The filter \mathbf{f} making a reference signal was set to be $\mathbf{f} = [1, 0, 0, 0]^T$. The Gaussian noises $n_i(t)$ ($i = 1, 2, 3, 4$) with their variances $\sigma_{n_i}^2$ were included in the outputs $y_i(t)$ at various SNR levels. The SNR was considered at the output of the system \mathbf{H} . As a measure of performance, we used the *multichannel intersymbol interference* (M_{ISI}) defined in [6].

For comparison, the SEA given by (21) and (22) was used. Fig. 2 (a), (b), and (c) show the results obtained by respectively first, second, and third iterations, using the EVA and the SEA. In each figure, the SNR levels were taken from 5 dB to 40 dB (see horizontal axis) and M_{ISI} 's shown in vertical axes were the average of the performance results obtained by 100 independent Monte Carlo runs. For each iteration, using three kinds of data samples, $t_1 = 1,000$ (case (i)), 2,500 (case (ii)), and 5,000 (case (iii)), the matrix \mathbf{C}_{yx} in the EVA, the matrix \mathbf{R} in the EVA and SEA, and the vector \mathbf{d}_l in the SEA were estimated by the following equations, respectively.

$$\begin{aligned} \mathbf{C}_{yx}(t) &:= \beta_1 \mathbf{C}_{yx}(t-1) + (1 - \beta_1) \{x^2(t) \mathbf{y}(t) \mathbf{y}^T(t) \\ &\quad - 2\tilde{v}_{x1}(t) x(t) \mathbf{y}^T(t) - \tilde{v}_{x2}(t) \mathbf{y}(t) \mathbf{y}^T(t)\}, \end{aligned} \quad (24)$$

$$\mathbf{R}(t) := \beta_2 \mathbf{R}(t-1) + (1 - \beta_2) \mathbf{y}(t) \mathbf{y}^T(t) \quad (25)$$

$$\begin{aligned} \mathbf{d}_l(t) &:= \beta_1 \mathbf{d}_l(t-1) + (1 - \beta_1) (z_l^3(t) \mathbf{y}(t) \\ &\quad - 3\tilde{v}_l(t) z_l(t) \mathbf{y}(t)), \quad l = 1, 2, 3, 4, \end{aligned} \quad (26)$$

where $\tilde{v}_{x1}(t)$, $\tilde{v}_{x2}(t)$, and $\tilde{v}_l(t)$ are the following moving averages defined by

$$\tilde{v}_{x1}(t) := \beta_2 \tilde{v}_{x1}(t-1) + (1 - \beta_2) x(t) \mathbf{y}(t), \quad (27)$$

$$\tilde{v}_{x2}(t) := \beta_2 \tilde{v}_{x2}(t-1) + (1 - \beta_2) x^2(t), \quad (28)$$

$$\tilde{v}_l(t) := \beta_2 \tilde{v}_l(t-1) + (1 - \beta_2) z_l^2(t), \quad l = 1, 2, 3, 4. \quad (29)$$

The parameter β_1 was set to be $(1 - 1/t_1)$ for each data sample set. The parameter β_2 was set to be 0.995 for first iteration and 0.998 for second and third iterations, in which this selection rule of β_2 was applied to all data sample set and was best rule for solving this BSS problem in our computer simulation. Here, when the equations (24) -(29) were used, the first

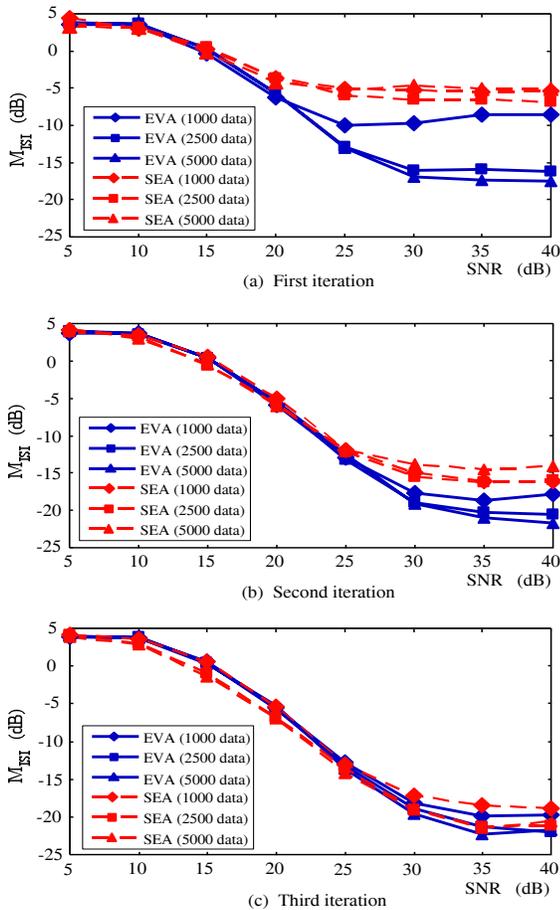


Fig. 2. The average performances of the EVA and the SEA with varying the SNR level, in the three cases of (i) 1000 samples, (ii) 2500 samples, (iii) 5000 samples.

iteration started with appropriate initial values, and the second and third iterations started with the initial values obtained from previous iteration. As a deflation method for the SEA, which is needed to separate all source signals, Gram-Schmidt type deflation method was adopted [2]. In the calculation of eigenvectors in the EVA, $C_{yx}^+ \mathbf{R}$ was used to obtain the eigenvectors with high accuracy.

From Fig. 2, it can be seen that the EVA gives better performances than the SEA. This means that if C_{yx} and \mathbf{R} are estimated with high accuracy, the EVA gives good results from *first iteration* (see, e.g., Fig. 2 (a)). On the other hand, the SEA does not provide such results, because the SEA needs deflation methods to obtain all source signals. However, this tendency is shown for the case where the SNR level is high and the number of iterations is small. As the number of iterations increases, this tendency disappears (see Fig. 2 (c)). Therefore we conclude that for the case where the SNR level is high and the number of iterations is small, the proposed EVA gives better performances, compared with the conventional algorithms which need deflation methods.

6. CONCLUSIONS

We have proposed an EVA for solving the BSS problem. By using reference signals, the proposed EVA is capable for separating source signals simultaneously from their mixtures. There exists a case that this point becomes a very attractive property for solving the BSS problems of MIMO systems, compared with the conventional algorithms which need deflation methods. Computer simulations have demonstrated the validity of the proposed EVA.

7. REFERENCES

- [1] A. Adib, E. Moreau, and D. Aboutajdine, "Source Separation Contrasts Using a Reference Signal," *IEEE Signal Processing Letters*, Vol. 11, No. 3, pp. 312-315, March 2004.
- [2] A. Hyvärinen, "Fast and robust fixed-point algorithms for independent component analysis," *IEEE Trans. Neural Networks*, vol. 10, no. 3, pp. 626-634, 1999.
- [3] Y. Inouye and K. Tanebe, "Super-exponential algorithms for multichannel blind deconvolution," *IEEE Trans. Signal Processing*, vol. 48, no. 3, pp. 881-888, Mar. 2000.
- [4] B. Jelonck and K. D. Kammeyer, "A closed-form solution to blind equalization," *Signal Processing*, 36(3), pp. 251-259, 1994.
- [5] B. Jelonck, D. Boss, and K. D. Kammeyer, "Generalized eigenvector algorithm for blind equalization," *Signal Processing*, 61(3), pp. 237-264, 1997.
- [6] M. Kawamoto, K. Kohno, and Y. Inouye, "Robust Super-Exponential Methods for Deflationary Blind Source Separation of Instantaneous Mixtures," *IEEE Transactions on Signal Processing*, Vol. 53, No. 5, pp. 1933-1937, May 2005.
- [7] M. Kawamoto, K. Kohno, and Y. Inouye, "Eigenvector Algorithms Using Reference Signals for Blind Source Separation of Instantaneous Mixtures," submitted to IS-CAS2006.
- [8] P. Lancaster and M. Tismenetsky, *The Theory of Matrices*, second edition, Academic Press, INC., 1985.
- [9] O. Shalvi and E. Weinstein, "New criteria for blind deconvolution of nonminimum phase system (channels)," *IEEE Trans. Inform. Theory*, vol. 36, no. 2, pp. 312-321, 1990.
- [10] O. Shalvi and E. Weinstein, "Super-exponential methods for blind deconvolution," *IEEE Trans. Inform. Theory*, vol. 39, no. 2, pp. 504-519, 1993.