ROBUST SUPER-EXPONENTIAL METHODS FOR BLIND DECONVOLUTION OF MIMO-FIR SYSTEMS WITH GAUSSIAN NOISE

Kiyotaka Kohno**, Yujiro Inouye*†*, and Mitsuru Kawamoto*‡

*,†Department of Electronic and Control Systems Engineering, Shimane University
1060 Nishikawatsu, Matsue, Shimane 690-8504, Japan
‡Information Technology Research Institute, National Institute of Advanced Industrial Science and Technology
Central 2, 1-1-1 Umezono, Tsukuba, Ibaraki 305-8568, Japan
email: *kohno@yonago-k.ac.jp, †inouye@riko.shimane-u.ac.jp, ‡m.kawamoto@aist.go.jp

ABSTRACT

The so called "super-exponential" methods (SEMs) are attractive methods for solving multichannel blind deconvolution problem. The conventional SEMs, however, have such a drawback that they are very sensitive to Gaussian noise. To overcome this drawback, the robust super-exponential methods (RSEMs) were proposed for single-input single-output infinite impulse response (SISO-IIR) channels and for multi-input multi-output (MIMO) static channels (instantaneous mixtures). While the conventional SEMs use the second- and higher-order cumulants of observations, the RSEMs use only the higher-order cumulants of observations. Since higher-order cumulants are insensitive to Gaussian noise, the RSEMs are robust to Gaussian noise. We proposed an RSEM extended to the case of MIMO-FIR channels (convolutive mixtures). To show the validity of the proposed RSEM, some simulation results are presented.

1. INTRODUCTION

The present paper deals with the multichannel blind deconvolution problem of finite-impulse response (FIR) channels of MIMO communication systems [1]. To solve this problem, the ideas of the super-exponential methods (SEMs) in [2]-[4] are used. Several researchers (e.g., [2]-[4]) have so far proposed some SEMs for solving independent component analysis (ICA), blind source separation (BSS) and blind channel equalization (BCE). One of the attractive properties of the SEMs is that the SEMs are computationally efficient and converge to a desired solution at a super-exponential rate. However, almost all the conventional SEMs have such a drawback that they are very sensitive to Gaussian noise, because the conventional SEMs utilize the second-order and the higher-order cumulants of observations.

To overcome the drawback, Kawamoto *et al.* proposed new SEMs for SISO-IIR channels [8] and for MIMO static channels (instantaneous mixtures) [7], and Kohno *et al.* proposed new SEMs for MIMO-IIR channels (convolutive mixtures) [9]. The proposed SEMs utilize only the higher-order cumulants of observations, and hence the proposed SEMs become robust to Gaussian noise, so that the proposed SEMs are referred to as *robust super-exponential methods* (RSEMs).

One may extend directly the previous idea of RSEMs to the case of MIMO channels (convolutive mixtures) when all the source signals are sub-Gaussian or super-Gaussian. However, this is not the case in general. The extension of the idea is not straightforward in the case when the source signals are of different types. Kohno *et al.* showed the perfect solution of blind deconvolution for MIMO-IIR channels in that case by making the length of the equalizer infinite (or considering an IIR equalizer).

The purpose of the present paper is to extend the idea for MIMO-IIR channels to the case of MIMO-FIR channels (convolutive mixtures) in the presence of Gaussian noise and to propose an RSEM to this case. We show an approximate solution of blind deconvolution for a MIMO-FIR channel by making the length of the equalizer very large, and we show that the proposed RSEM treats such general cases as some of the source signals are sub-Gaussian and the remainder are super-Gaussian. Simulation results are presented to show the effectiveness of the proposed RSEM.

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 $\boldsymbol{C}^{m\times n}$ denote the set of all $m\times n$ matrices with complex components. The superscripts T, $*$, H and $\dagger$ denote, respectively, the transpose, the complex conjugate, the complex conjugate transpose (Hermitian) and the (Moore-Penrose) pseudoinverse operations of a matrix or a linear operator. The symbols Ker $\boldsymbol{A}$ and Im $\boldsymbol{A}$ denote the kernel and the image of matrix $\boldsymbol{A}$, respectively. The superscript $\perp$ denotes the orthogonal complement of a subspace. Let $i=\overline{1,n}$ stand for $i=1,2,\cdots,n$.

2. PROBLEM FORMULATION

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

$$\boldsymbol{y}(t)=\sum_{k=0}^{K-1}\boldsymbol{H}^{(k)}\boldsymbol{s}(t-k)+\boldsymbol{n}(t),\quad t\in Z, \tag{1}$$

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 number K denotes its length. The transfer function of the channel is defined by

$$\boldsymbol{H}(z)=\sum_{k=0}^{K-1}\boldsymbol{H}^{(k)}z^{-k},\quad z\in C. \tag{2}$$

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

$$\boldsymbol{z}(t) \quad = \quad \sum_{k=0}^{L-1}\boldsymbol{W}^{(k)}\boldsymbol{y}(t-k)$$

$$= \sum_{k=0}^{K+L-2} \boldsymbol{G}^{(k)}\boldsymbol{s}(t-k) + \sum_{k=0}^{L-1} \boldsymbol{W}^{(k)}\boldsymbol{n}(t-k), \quad (3)$$

where the number L is the length of the deconvolver, and $\{\boldsymbol{G}^{(k)}\}$ is the impulse response of the cascade system of the unknown system $\boldsymbol{H}(z)$ and the deconvolver $\boldsymbol{W}(z)$ defined by

$$\boldsymbol{G}^{(k)} := \sum_{\tau=0}^{L-1} \boldsymbol{W}^{(\tau)}\boldsymbol{H}^{(k-\tau)}, \quad k = \overline{0, K+L-2}. \quad (4)$$

Here we should note that the length of the cascade system is $K+L-1$.

The objective of multichannel blind deconvolution is to construct a deconvolver that recovers the original source signals only from the measurements of the corresponding outputs.

We put the following assumptions on the channel and the source signals.

A1) The channel system $\boldsymbol{H}(z)$ is an FIR system which is equalizable. The equalizability condition of $\boldsymbol{H}(z)$ means that $\boldsymbol{H}(z)$ has no zero on z-plane except for the origin $z=0$ [10], which is equivalent to

$$\text{rank } \boldsymbol{H}(\lambda) = n \quad \text{for all nonzero } \lambda \in C. \quad (5)$$

A2) The input sequence $\{\boldsymbol{s}(t)\}$ is a complex, zero-mean 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 independently and identically distributed (i.i.d.) process with a variance $\sigma_i^2 \neq 0$ and a fourth-order cumulant $\gamma_i \neq 0$. Moreover, each element process $\{s_i(t)\}$ has nonzero $(p+q+1)$ st-order cumulants κ_i defined as

$$\kappa_i = \text{cum}\{\underbrace{s_i(t), \cdots, s_i(t)}_{p}, \underbrace{s_i^*(t), \cdots, s_i^*(t)}_{q+1}\} \neq 0, \quad (6)$$

where p and q are nonnegative integers such that $p+q \geq 2$.

A3) The deconvolver (or equalizer) $\boldsymbol{W}(z)$ is an FIR system of length $L \geq L_o(H)$, where $L_o(H)$ is the minimum length of its deconvolvers which attain perfect equalization of $\boldsymbol{H}(z)$ in the noiseless case (see Remark 1 below).

A4) The noise sequence $\{\boldsymbol{n}(t)\}$ is a zero-mean, Gaussian vector stationary process.

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

Remark 1. The assumption on A1) means that the unknown system $\boldsymbol{H}(z)$ has less inputs than outputs, i.e., $n \leq m$ and that there exists an FIR left inverse of $\boldsymbol{H}(z)$. Moreover, if $\boldsymbol{H}(z)$ has no zero on the z-plane, that is, it is irreducible in the sense that it satisfies

$$\text{rank } \boldsymbol{H}(\lambda) = n \quad \text{for all } \lambda \in C, \quad (7)$$

then there exists an equalizer $\boldsymbol{W}(z)$ of length $L \leq n(K-1)$, where K is the length of the channel [10]. Besides, it is shown in [4] that there exists generically (or except for pathological cases) an equalizer $\boldsymbol{W}(z)$ of length $L = \lceil \frac{n(K-1)}{m-n} \rceil$, where $\lceil x \rceil$ stands for the smallest integer that is greater than or equal to x.

In a vector form, (4) can be written as

$$\tilde{\boldsymbol{g}}_i = \tilde{\boldsymbol{H}}\tilde{\boldsymbol{w}}_i, \quad i = \overline{1,n}, \quad (8)$$

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

$$\tilde{\boldsymbol{g}}_i := [\boldsymbol{g}_{i1}^T, \boldsymbol{g}_{i2}^T, \cdots, \boldsymbol{g}_{in}^T]^T, \quad (9)$$

$$\boldsymbol{g}_{ij} := [g_{ij}(0), g_{ij}(1), \cdots, g_{ij}(K+L-2)]^T, \quad (10)$$

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

$$\tilde{\boldsymbol{w}}_i := [\boldsymbol{w}_{i1}^T, \boldsymbol{w}_{i2}^T, \cdots, \boldsymbol{w}_{im}^T]^T \in \boldsymbol{C}^{mL}, \quad (11)$$

$$\boldsymbol{w}_{ij} := \left[w_{ij}{}^{(0)}, w_{ij}{}^{(1)}, \cdots, w_{ij}{}^{(L-1)}\right]^T \in \boldsymbol{C}^L, \quad (12)$$

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

$$\tilde{\boldsymbol{H}} := \begin{bmatrix} \boldsymbol{H}_{11} & \boldsymbol{H}_{12} & \cdots & \boldsymbol{H}_{1m} \\ \boldsymbol{H}_{21} & \boldsymbol{H}_{22} & \cdots & \boldsymbol{H}_{2m} \\ \vdots & \vdots & \vdots & \vdots \\ \boldsymbol{H}_{n1} & \boldsymbol{H}_{n2} & \cdots & \boldsymbol{H}_{nm} \end{bmatrix}, \quad (13)$$

whose (i,j)th block element $\boldsymbol{H}_{ij}$ is the matrix (of L columns and $K+L-1$ rows) with the (l,r)th element $[\boldsymbol{H}_{ij}]_{lr}$ defined by

$$[\boldsymbol{H}_{ij}]_{lr} := h_{ji}(l-r),$$
$$l = 0,1,\cdots,K+L-2, \quad r = 0,1,\cdots,L-1. \quad (14)$$

Here $h_{ij}(n) = 0$ for $n < 0$.

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

$$[\tilde{\boldsymbol{g}}_1, \cdots, \tilde{\boldsymbol{g}}_n] = \tilde{\boldsymbol{H}}[\tilde{\boldsymbol{w}}_1, \cdots, \tilde{\boldsymbol{w}}_n] = [\tilde{\boldsymbol{\delta}}_1, \cdots, \tilde{\boldsymbol{\delta}}_n]\boldsymbol{P}, \quad (15)$$

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

$$\tilde{\boldsymbol{\delta}}_i := [\boldsymbol{\delta}_{i1}^T, \boldsymbol{\delta}_{i2}^T, \ldots, \boldsymbol{\delta}_{in}^T]^T, \quad (16)$$

$$\boldsymbol{\delta}_{ij} := \begin{cases} \hat{\boldsymbol{\delta}}_i, & \text{if } i = j, \\ (\underbrace{0,0,\cdots,0}_{K+L-1})^T, & \text{otherwise.} \end{cases} \quad (17)$$

Here, $\hat{\boldsymbol{\delta}}_i$ is the column vector (of $K+L-1$ elements) whose rth element $\hat{\delta}_i(r)$ given by

$$\hat{\delta}_i(r) = d_i \delta(r - k_i), \quad (18)$$

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 a integer standing for a time shift.

3. ROBUST SUPER-EXPONENTIAL METHODS

3.1 Two-step iterative procedure for vector $\tilde{\boldsymbol{g}}_i$

To find solutions in (15), the following two-step iterative procedure with respect to the elements g_{ij}, $j = \overline{1,n}$ of the vector $\tilde{\boldsymbol{g}}_i$ is used:

$$g_{ij}^{[1]}(k) = \frac{\kappa_j}{a_j(k)\gamma_j}(g_{ij}(k))^p(g_{ij}^*(k))^q, \quad j = \overline{1,n}, \quad (19)$$

$$g_{ij}^{[2]}(k) = \frac{g_{ij}^{[1]}(k)}{\sqrt{\sigma_{z_i}^2}}, \quad j = \overline{1,n}, \quad (20)$$

where $g_{ij}(k)$ in the right-hand side of (19) is an element of $\tilde{\boldsymbol{g}}_i$ before the iteration, $(\cdot)^{[1]}$ and $(\cdot)^{[2]}$ stand for the results of the first step and the second step per iteration, p and q are nonnegative integers such that $p+q \geq 2$, $a_j(k)$ denotes a positive number (in Subsection 3.2, it will be shown how we choose the values of $a_j(k)$'s), γ_j denotes the fourth-order cumulant of $s_j(t)$, that is, γ_j is equal to κ_j in case of p = 2 and q = 1, and $\sigma_{z_i}^2$ denotes the variance of the output signal $z_i(t)$. Equation (19) is derived by replacing σ_j^2 of (26) in [3] with $a_j(k)\gamma_j$, where σ_j^2 denotes the second-order cumulant of $s_j(t)$, and (20) is used to normalize $g_{ij}^{[1]}$ obtained by (19).

Here it should be noted that in the conventional two-step procedures (e.g., [2]-[4]), the denominator of the right-hand side of (19) was set to be 1 or the variance of $s_j(t)$, whereas we consider the fourth-order cumulant of $s_j(t)$, i.e., γ_j.

Let $g_{ij}^{(l)}(k)$ denote the value obtained in the lth cycle of the iterations of two steps (19) and (20). The important fact of the two-step procedure is that the n values $g_{ij}^{(l)}(k)$ ($j=\overline{1,n}$) converge to zero except for only one of the values as the iteration number l approaches infinity, that is, $l\to\infty$. The magnitude of the remaining one converges to a positive constant. This will be shown in the following theorem.

Theorem 1. Let $g_{ij}^{(0)}(k)$ be an initial value for iterations of two steps (19) and (20) for each $j=\overline{1,n}$ and $k\in Z$. Let $\alpha_j(k)$ be a non-negative scalar defined as

$$\alpha_j(k)=\left|\frac{\kappa_j}{a_j(k)\gamma_j}\right|^{\frac{1}{p+q-1}}. \tag{21}$$

Let j_i and k_i be $(j_i,k_i)=\arg\max_{(j,k)}\alpha_j(k)|g_{ij}^{(0)}(k)|$. Suppose the index j_i and k_i are unique, that is, $\alpha_{j_i}(k_i)|g_{ij_i}^{(0)}(k_i)|>\alpha_j(k)|g_{ij}^{(0)}(k)|$ for any other j and k. Then as $l\to\infty$, it follows

$$\lim_{l\to\infty}|g_{ij}^{(l)}(k)|=\begin{cases}0, & \text{for } j\neq j_i,\quad k\neq k_i,\\ \tilde{c}_j(k)\neq 0, & \text{for } j=j_i,\quad k=k_i,\end{cases} \tag{22}$$

where $\tilde{c}_j(k)$ is a scalar positive constant.

Since Theorem 1 may be proven by using the similar way as in [7], the proof of Theorem 1 is omitted for page limit.

For notational simplicity, we confine ourselves to the case $p=2$ and $q=1$ (which gives a solution in terms of fourth-order cumulants), although our results are expandable to a general (p,q) case (higher order cumulant case).

We turn to the two-step procedure (19) and (20) with $p=1$, $q=2$ and $\kappa_j=\gamma_j$ ($j=\overline{1,n}$). It becomes

$$g_{ij}^{[1]}(k)=\frac{1}{a_j(k)}(g_{ij}(k))^2(g_{ij}^*(k)),\qquad j=\overline{1,n}, \tag{23}$$

$$g_{ij}^{[2]}(k)=\frac{g_{ij}^{[1]}(k)}{\sqrt{\sigma_{z_i}^2}},\qquad j=\overline{1,n}. \tag{24}$$

3.2 Two-step iterative procedure for vector $\tilde{\boldsymbol{w}}_i$

Since the parameters $g_{ij}(k)$'s involve implicitly the unknown parameters $h_{ij}(k)$'s, the two-step procedure cannot be handled directly. Therefore, by solving the following weighted least squares problem, we derive an algorithm with respect to w_{ij}'s so that the two steps (23) and (24) can be handled directly.

$$\min_{\tilde{\boldsymbol{w}}_i}(\tilde{\boldsymbol{H}}\tilde{\boldsymbol{w}}_i-\tilde{\boldsymbol{g}}_i)^H\tilde{\Lambda}(\tilde{\boldsymbol{H}}\tilde{\boldsymbol{w}}_i-\tilde{\boldsymbol{g}}_i),\qquad i=\overline{1,n}. \tag{25}$$

Here, $\tilde{\Lambda}$ is a diagonal matrix with positive diagonal elements. The solutions are known to be given by

$$\tilde{\boldsymbol{w}}_i=(\tilde{\boldsymbol{H}}^H\tilde{\Lambda}\tilde{\boldsymbol{H}})^{\dagger}\tilde{\boldsymbol{H}}^H\tilde{\Lambda}\tilde{\boldsymbol{g}}_i,\qquad i=\overline{1,n}. \tag{26}$$

In the conventional methods [2]-[4], the positive diagonal elements of $\tilde{\Lambda}$ are set to be 1 or the variances of the source signals. This means that $\tilde{\boldsymbol{H}}^H\ \tilde{\Lambda}\ \tilde{\boldsymbol{H}}$ is calculated by the second-order statistics of the observed signal $\boldsymbol{y}(t)$. We are convinced that this is the reason why the conventional methods are sensitive to Gaussian noise.

In what follows, we shall show that the weighted least squares approach in (25) can be applied to a set of fourth-order cumulants of the observed signals $y_i(t)$ ($i=\overline{1,m}$), if we choose appropriately a diagonal matrix $\tilde{\Lambda}$ in (25). To this end, we introduce fourth-order cumulants matrices of m-vector random process $\{\boldsymbol{y}(t)\}$, which constitute a set of $m\times m$ block matrices $\tilde{\boldsymbol{C}}_{\boldsymbol{y}_{i,j,l}}^{(4)}$, whose elements are defined by

$$\left[\tilde{\boldsymbol{C}}_{\boldsymbol{y}_{i,j,l}}^{(4)}\right]_{[p,q]_{l_1l_2}}=\text{cum}\{y_q(t-l_2),y_p^*(t-l_1),y_j(t-l),y_i^*(t-l)\},$$
$$p,q,i,j=\overline{1,m},\quad l_1,l_2=\overline{0,L-1},\quad l=\overline{0,K+L-2}, \tag{27}$$

where $[\cdot]_{[p,q]_{l_1l_2}}$ denotes the (l_1,l_2)th element of the (p,q)th block matrix of the matrix $\tilde{\boldsymbol{C}}_{\boldsymbol{y}_{i,j,l}}^{(4)}$. Then, we consider an $m\times m$ block matrix $\tilde{\boldsymbol{C}}$ expressed by

$$\tilde{\boldsymbol{C}}=\sum_{i,j=1}^{m}\sum_{l=0}^{K+L-2}\beta_{ij}\tilde{\boldsymbol{C}}_{\boldsymbol{y}_{i,j,l}}^{(4)}, \tag{28}$$

where β_{ij}'s are either 1 or 0, which represent *design parameters*. It is shown by a simple calculation that (28) becomes

$$\tilde{\boldsymbol{C}}=\tilde{\boldsymbol{H}}^H\tilde{\Sigma}\tilde{\boldsymbol{H}}, \tag{29}$$

where $\tilde{\Sigma}$ is a diagonal matrix defined by

$$\tilde{\Sigma}:=\text{diag}\{\Sigma_1,\Sigma_2,\cdots,\Sigma_n\}, \tag{30}$$

$$\Sigma_r:=\text{diag}\{\gamma_r\tilde{a}_r(0),\gamma_r\tilde{a}_r(1),\cdots,\gamma_r\tilde{a}_r(K+L-2)\},r=\overline{1,n}, \tag{31}$$

$$\tilde{a}_r(k):=\sum_{i,j=1}^{m}\sum_{l=0}^{K+L-2}\beta_{ij}h_{ir}(k-l)h_{jr}^*(k-l),k=\overline{0,K+L-2}, \tag{32}$$

where $\text{diag}\{\cdots\}$ denotes a diagonal matrix with the diagonal elements built from its arguments.

Here we note that the diagonal matrix $\tilde{\Sigma}$ is not positive semi-definite but the diagonal matrix $\tilde{\Lambda}$ defined by

$$\tilde{\Lambda}:=\text{diag}\{\Lambda_1,\Lambda_2,\cdots,\Lambda_n\}, \tag{33}$$

$$\Lambda_r=\text{diag}\{|\gamma_r\tilde{a}_r(0)|,|\gamma_r\tilde{a}_r(1)|,\cdots,|\gamma_r\tilde{a}_r(K+L-2)|\} \tag{34}$$

is positive semi-definite. It is clear from the definitions (30) and (33) that there exists a sign matrix $\dot{\boldsymbol{I}}$ such that $\tilde{\Lambda}=\tilde{\Sigma}\dot{\boldsymbol{I}}$, where the sign matrix $\dot{\boldsymbol{I}}$ is defined as a diagonal matrix whose diagonal elements are either $+1$ or -1.

In (28), let $\beta_{ij}=1$ for $i=j$ and $\beta_{ij}=0$ for $i\neq j$, then $\tilde{a}_r(k)$'s of the diagonal elements of $\tilde{\Sigma}$ become

$$\tilde{a}_r(k)=\sum_{i=1}^{m}\sum_{l=0}^{K+L-2}|h_{ir}(k-l)|^2\geq 0,r=\overline{1,n},k=\overline{0,K+L-2}. \tag{35}$$

Therefore, all the diagonal elements of $\tilde{\Sigma}$ and $\tilde{\Lambda}$ are nonzero except for pathological cases.

Theorem 2. Let

$$\tilde{\boldsymbol{w}}_i(\tilde{\Lambda}):=(\tilde{\boldsymbol{H}}^H\tilde{\Lambda}\tilde{\boldsymbol{H}})^{\dagger}\tilde{\boldsymbol{H}}^H\tilde{\Lambda}\tilde{\boldsymbol{g}}_i,\quad i=\overline{1,n}, \tag{36}$$

and

$$\tilde{\boldsymbol{w}}_i(\tilde{\Sigma}):=(\tilde{\boldsymbol{H}}^H\tilde{\Sigma}\tilde{\boldsymbol{H}})^{\dagger}\tilde{\boldsymbol{H}}^H\tilde{\Sigma}\tilde{\boldsymbol{g}}_i,\quad i=\overline{1,n}, \tag{37}$$

where $\tilde{\Lambda}$ and $\tilde{\Sigma}$ are diagonal invertible matrices, and $\tilde{\boldsymbol{H}}$ and $\tilde{\boldsymbol{g}}_i$'s are defined by (13) and (9) (along with (10)), respectively. Then

(a) If $\tilde{\Lambda}=\tilde{\Sigma}$ or $\tilde{\Lambda}=-\tilde{\Sigma}$, then

$$\tilde{\boldsymbol{w}}_i(\tilde{\Lambda})=\tilde{\boldsymbol{w}}_i(\tilde{\Sigma}),\quad i=\overline{1,n}. \tag{38}$$

(b) If $\boldsymbol{H}(z)$ satisfies Assumption A1) and the length L is infinite (i.e. $L=+\infty$), then

$$\tilde{\boldsymbol{w}}_i(\tilde{\Lambda})=\tilde{\boldsymbol{w}}_i(\tilde{\Sigma}),\quad i=\overline{1,n}. \tag{39}$$

To prove the statement (b) of Theorem 2, we have recourse to the following lemma.

Lemma 1. Let $\boldsymbol{A}$ and $\boldsymbol{B}$ are matrices of infinite dimension. Then the following facts hold true.
(a) If Ker $\boldsymbol{A} = \{0\}$, then $\boldsymbol{A}^{\dagger}\boldsymbol{A} = \boldsymbol{I}$.
(b) If Im $\boldsymbol{A}$ = (Ker $\boldsymbol{B})^{\perp}$, then $(\boldsymbol{B}\boldsymbol{A})^{\dagger} = \boldsymbol{A}^{\dagger}\boldsymbol{B}^{\dagger}$. The proof of Lemma 1 is omitted for page limit.

Proof of Theorem 2.
(a) If $\tilde{\Lambda} = \tilde{\Sigma}$, then (38) follows immediately from (36) and (37). If $\tilde{\Lambda} = -\tilde{\Sigma}$, then (38) also follows immediately from (36) and (37).
(b) Suppose $L = +\infty$. Let

$$\tilde{\boldsymbol{y}} := \tilde{\boldsymbol{H}}^{T}\tilde{\boldsymbol{s}}, \tag{40}$$

where we use the same notation as in (9) and (10) for the elements of $\tilde{\boldsymbol{y}}$ and $\tilde{\boldsymbol{s}}$, and thus they are defined as

$$\tilde{\boldsymbol{y}} := \left[\boldsymbol{y}_1^T, \boldsymbol{y}_2^T, \cdots, \boldsymbol{y}_m^T\right]^T, \tag{41}$$

$$\boldsymbol{y}_i := [y_i(0), y_i(1), \cdots]^T,\ i = \overline{1,m}, \tag{42}$$

$$\tilde{\boldsymbol{s}} := \left[\boldsymbol{s}_1^T, \boldsymbol{s}_2^T, \cdots, \boldsymbol{s}_n^T\right]^T, \tag{43}$$

$$\boldsymbol{s}_i := [s_i(0), s_i(1), \cdots]^T,\ i = \overline{1,n}. \tag{44}$$

In the time domain, (40) is equivalent to

$$\boldsymbol{y}(t) = \sum_{k=0}^{K-1} \boldsymbol{H}^{(k)}\boldsymbol{s}(t-k), \quad t = 0,1,2,\cdots. \tag{45}$$

Therefore, suppose $\tilde{\boldsymbol{y}} = 0$ which is equivalent to $\boldsymbol{y}(t) = 0$ for all $t = 0,1,2,\cdots$. If follows from (45) and A1) that $\boldsymbol{s}(t) = 0$ for all $t = 0,1,2,\cdots$, which is equivalent to $\tilde{\boldsymbol{s}} = 0$. Taking the complex conjugates of elements on the both side of (40), we obtain

$$\text{Ker } \tilde{\boldsymbol{H}}^{H} = \{0\}. \tag{46}$$

On the other hand, it is well known (see [6])

$$\text{Ker } \tilde{\boldsymbol{H}}^{H} = (\text{Im } \tilde{\boldsymbol{H}})^{\perp}. \tag{47}$$

Since $\tilde{\Lambda}$ and $\tilde{\Sigma}$ are invertible, it follows from (46)

$$\text{Ker } \tilde{\boldsymbol{H}}^{H}\tilde{\Lambda} = \{0\} \quad \text{and} \quad \text{Ker } \tilde{\boldsymbol{H}}^{H}\tilde{\Sigma} = \{0\}. \tag{48}$$

Using the statement (b) in Lemma 1, (36) becomes

$$\tilde{\boldsymbol{w}}_i(\tilde{\Lambda}) = \tilde{\boldsymbol{H}}^{\dagger}(\tilde{\boldsymbol{H}}^{H}\tilde{\Lambda})^{\dagger}\tilde{\boldsymbol{H}}^{H}\tilde{\Lambda}\tilde{\boldsymbol{g}}_i, \quad i = \overline{1,n}, \tag{49}$$

and (37) becomes

$$\tilde{\boldsymbol{w}}_i(\tilde{\Sigma}) = \tilde{\boldsymbol{H}}^{\dagger}(\tilde{\boldsymbol{H}}^{H}\tilde{\Sigma})^{\dagger}\tilde{\boldsymbol{H}}^{H}\tilde{\Sigma}\tilde{\boldsymbol{g}}_i, \quad i = \overline{1,n}. \tag{50}$$

Using the statement (a) in Lemma 1, (49) becomes

$$\tilde{\boldsymbol{w}}_i(\tilde{\Lambda}) = \tilde{\boldsymbol{H}}^{\dagger}\tilde{\boldsymbol{g}}_i, \quad i = \overline{1,n}, \tag{51}$$

and (50) becomes

$$\tilde{\boldsymbol{w}}_i(\tilde{\Sigma}) = \tilde{\boldsymbol{H}}^{\dagger}\tilde{\boldsymbol{g}}_i, \quad i = \overline{1,n}. \tag{52}$$

Therefore, we obtain equalities

$$\tilde{\boldsymbol{w}}_i(\tilde{\Lambda}) = \tilde{\boldsymbol{w}}_i(\tilde{\Sigma}), \quad i = \overline{1,n}. \tag{53}$$

This completes the proof of Theorem 2.

Remark 2. Based on the statement (b) in Theorem 2, if the parameter L of the deconvolver in (12) is chosen to be enough large positive values (let us say, $L \simeq +\infty$), then we have approximate relations

$$\tilde{\boldsymbol{w}}_i(\tilde{\Lambda}) \simeq \tilde{\boldsymbol{w}}_i(\tilde{\Sigma}), \quad i = \overline{1,n}. \tag{54}$$

Therefore, the proposed method can be applied to the case when the signs of the fourth-order cumulants γ_i $(i = \overline{1,n})$ are different (let us say, we can treat sub-Gaussian and super-Gaussian signals as the elements of source vector $\boldsymbol{s}(t)$).

For the time being, in the present paper, we consider (28) with β_{ij} = 1 for $i = j$ and β_{ij} = 0 for $i \neq j$. As for $\tilde{\boldsymbol{H}}^{H}\tilde{\Sigma}\tilde{\boldsymbol{H}}$, (28) can be estimated recursively by the fourth-order cumulants block matrices of $\boldsymbol{y}(t)$ using moving averages as shown in [8].

Moreover, as for $\tilde{\boldsymbol{H}}^{H}\tilde{\Sigma}\tilde{\boldsymbol{g}}_i$, by using (23) with $a_j(k)$ = $\tilde{a}_j(k)$ in (35) and the similar way as in [3], it can be calculated by

$$\tilde{\boldsymbol{D}}_i := [\boldsymbol{d}_{i1}^T, \boldsymbol{d}_{i2}^T, \cdots, \boldsymbol{d}_{im}^T]^T, \tag{55}$$

$$[\boldsymbol{d}_{ij}]_l := \text{cum}\{z_i(t), z_i(t), z_i^*(t), y_j^*(t-l)\}, \tag{56}$$

$$i = \overline{1,n}, \quad j = \overline{1,m}, \quad l = 0,1,\cdots,L-1.$$

Then (26) can be expressed as

$$\tilde{\boldsymbol{w}}_i^{[1]} := \tilde{\boldsymbol{C}}^{\dagger}\tilde{\boldsymbol{D}}_i, \quad i = \overline{1,n}, \tag{57}$$

Since the second step (24) is a normalization of $\tilde{\boldsymbol{g}}_i$, it is easily shown that the second step reduces to

$$\tilde{\boldsymbol{w}}_i^{[2]} = \tilde{\boldsymbol{w}}_i^{[1]}/\sqrt{\sigma_{z_i}^2} \quad i = \overline{1,n}. \tag{58}$$

Therefore, (57) and (58) are our proposed two steps to modify $\tilde{\boldsymbol{w}}_i$, which constitutes one cycle of iterations in the super-exponential method [2]-[4]. Then since the right-hand side of (57) consists of only fourth-order cumulants, the modification of $\tilde{\boldsymbol{w}}_i$ is not affected by Gaussian noise. This comes from the fact that higher-order cumulants are insensitive to additive (even colored) Gaussian noise. This is a *novel key point* of our proposed super-exponential method, from which the proposed method is referred also to as a *robust super-exponential method* (RSEM).

4. SIMULATION RESULTS

To demonstrate the validity of the proposed RSEM, many computer simulations were conducted. The deflation method [3] was employed for the proposed RSEM in our simulations. Two important simulation results are shown in this section. We considered a MIMO channel with two inputs and three outputs, and assumed that the length of channel is three (K = 3), that is $\boldsymbol{H}^{(k)}$s in (1) were set to be

$$\boldsymbol{H}(z) = \sum_{k=0}^{2}\boldsymbol{H}^{(k)}z^{-k} =$$

$$\begin{bmatrix} 1.00+0.15z^{-1}+0.10z^{-2} & 0.65+0.25z^{-1}+0.15z^{-2} \\ 0.50-0.10z^{-1}+0.20z^{-2} & 1.00+0.25z^{-1}+0.10z^{-2} \\ 0.60+0.10z^{-1}+0.40z^{-2} & 0.10+0.20z^{-1}+0.10z^{-2} \end{bmatrix} \tag{59}$$

The length of the deconvolver was chosen to be four ($L = 4$) and eight ($L = 8$). We set the values of the tap coefficients to be zero except for $w_{11}^{(3)} = w_{22}^{(3)}$ = 1 in case of $L = 4$ and $w_{11}^{(5)} = w_{22}^{(5)}$ = 1 in case of $L = 8$. Two source signals $s_1(t)$ and $s_2(t)$ were a sub-Gaussian signal and a super-Gaussian signal, in which the sub-Gaussian signal takes one of two values, -1 and 1 with same probability 1/2, and the super-Gaussian signal takes $-\sqrt{5}$, $\sqrt{5}$ and 0 with probabilities 1/10, 1/10 and 4/5, respectively. The parameter p and q in (6) were set to be p = 2 and q = 1, that is, κ_j (j = 1,2) in (19) were the fourth-order cumulants of the source signals. The values of the fourth-order cumulants of the sub-Gaussian and super-Gaussian signal were -2 and +2, respectively. Three independent Gaussian noises (with identical variance σ_n^2) were added to the three outputs $y_i(t)$'s at various SNR levels. The SNR is, for convenience, defined as SNR := 10 $\log_{10}(\sigma_{s_i}^2/\sigma_n^2)$, where $\sigma_{s_i}^2$'s are the variances of $s_i(t)$'s and are equal to 1.

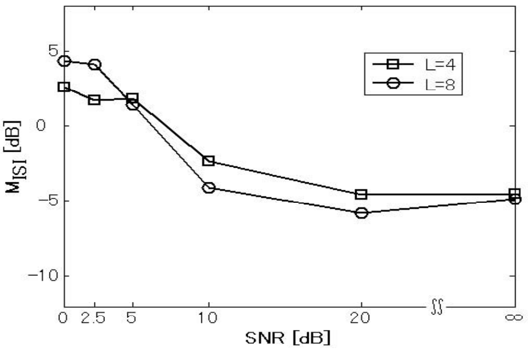

Figure 1: The performances of the proposed RSEM for the length of deconvolver $L = 4$ and $L = 8$.

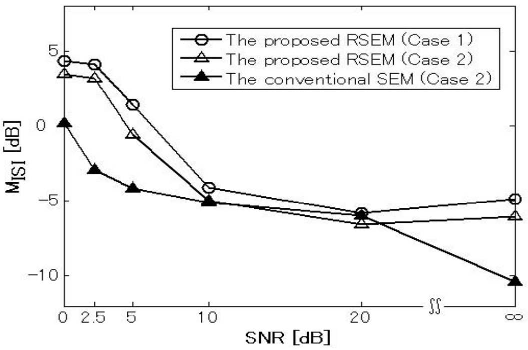

Figure 2: The performances for the proposed RSEM and the conventional SEM.

As a measure of performance, we used the *multichannel intersymbol interference* (M_{ISI}) [3],[7]. The value of M_{ISI} becomes $-\infty$, if $\tilde{\boldsymbol{g}}_l$'s in (8) are obtained, and hence a minus large value of M_{ISI} indicates the proximity to the desired solution.

First we consider an effect of the length of the deconvolver (L). Fig. 1 shows the results of performances of the proposed RSEM for the length of deconvolver $L = 4$ and $L = 8$ when the SNR levels were respectively taken to be 0[dB] ($\sigma_n^2 = 1$), 2.5[dB], 5[dB], 10[dB], 20[dB] and ∞ [dB] ($\sigma_n^2 = 0$), in which each M_{ISI} shown in Fig. 1 was the average of the performance results obtained by 10 independent Monte Carlo runs. In each Monte Carlo run, $\tilde{\boldsymbol{C}}$ and $\tilde{\boldsymbol{D}}_i$ were estimated by 10,000 data samples.

It can be seen from Fig. 1 that the performance of the proposed RSEM is better as the number of the length of the deconvolver increases at SNR greater than 5 dB. This implies that the performance of the RSEM depends on the length of the deconvolver and that the mixed type source signals of the sub-Gaussian and the super-Gaussian are deconvolved better as the length of the deconvolver increases. This fact is shown in Theorem 2.

Secondly we compare the proposed RSEM with the conventional method (SEM) [3]. Fig. 2 shows the results of performances for both the proposed RSEM and the conventional SEM by the same SNR levels and the same number of independent Monte Carlo runs as in Fig. 1. In each Monte Carlo run, $\tilde{\boldsymbol{C}}$ and $\tilde{\boldsymbol{D}}_i$ were estimated by data samples in the following two cases; (Case 1) 10,000 data and (Case 2) 30,000 data. The length of the deconvolver was chosen to be eight ($L = 8$).

It can be seen from Fig. 2 that the proposed RSEM shows better performance than the conventional SEM in Case 2 at SNR level is in the vicinity of 20 dB, and as the number of data samples which are needed to estimate the cumulants increases, the proposed RSEM shows much better performance.

5. CONCLUSIONS

We proposed an RSEM for deconvolving blindly MIMO-FIR channels in the presence of Gaussian noise. It can treat such general cases as some of the source signals are sub-Gaussian and the remainder are super-Gaussian. It was shown from the simulation results that the proposed RSEM is robust to Gaussian noise and can successfully solve the multichannel blind deconvolution problem.

ACKNOWLEDGMENT

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

REFERENCES

[1] Erik G. Larsson and Petre Stoica, *Space-Time Block Coding for Wireless Communications*, Cambridge University Press, 2003.

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

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

[5] P. Lancaster and M. Tismenetsky, *The Theory of Matrices*, second edition, Academic Press, INC., 1985.

[6] C. A. Desorer and B. H. Whalen, “A note on pseudoinverses,” *J. Soc. Indust. Appl. Math.*, vol. 11, no. 2, pp. 442-447, Jun. 1963.

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

[8] M. Kawamoto, M. Ohata, K. Kohno, Y. Inouye and Asoke K. Nandi, “Robust super-exponential methods for blind equalization in the presence of Gaussian noise,” *IEEE Trans. Circuits and Systems*-II, vol. 52, no. 10, pp. 651-655, Oct. 2005.

[9] K. Kohno, Y. Inouye and M. Kawamoto, “Robust super-exponential methods for blind deconvolution of MIMO-IIR systems with Gaussian noise,” *Proc. of ISCAS 2006*, Island of Kos, Greese, May 21-24, 2006, pp. 3598-3601.

[10] Y. Inouye and R-W. Liu, “A system-theoretic foundation for blind equalization of an FIR MIMO channel system,” *IEEE Trans. Circuits and Systems*-I, *Fundam. Theory Appl.* vol. 49, no. 4, pp. 425-436, Apr. 2002.