[28] ——, "Sampling—50 years after Shannon," *Proc. IEEE*, vol. 88, no. 4, pp. 569–589, Apr. 2000.
[29] M. Unser and A. Aldroubi, "A general sampling theory for nonideal acquisition devices," *IEEE Trans. Signal Process.*, vol. 42, no. 11, pp. 2915–2925, Nov. 1994.
[30] G. G. Walter, "A sampling theorem for wavelet subspaces," *IEEE Trans. Inf. Theory*, vol. 38, no. 2, pp. 881–884, Mar. 1992.
[31] E. T. Whittaker, "On the functions which are represented by the expansion of interpolating theory," in *Proc. R. Soc. Edinburgh*, vol. 35, 1915, pp. 181–194.
[32] X. G. Xia and Z. Zhang, "On sampling theorem, wavelets, and wavelet transforms," *IEEE Trans. Signal Process.*, vol. 41, no. 12, pp. 2525–3524, Dec. 1993.

Robust Super-Exponential Methods for Deflationary Blind Source Separation of Instantaneous Mixtures

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

***Abstract*—The so-called "super-exponential" methods (SEMs) are attractive methods for solving blind signal processing problems. The conventional SEMs, however, have such a drawback that they are very sensitive to Gaussian noise. To overcome this drawback, we propose a new SEM. While the conventional SEMs use the second- and higher order cumulants of observations, the proposed SEM uses only the higher order cumulants of observations. Since higher order cumulants are insensitive to Gaussian noise, the proposed SEM is robust to Gaussian noise, which is referred to as a robust super-exponential method (RSEM). To show the validity of the proposed RSEM, some simulation results are presented.**

***Index Terms*—Blind source separation, deflationary approach, Gaussian noise, instantaneous mixtures, super-exponential methods.**

I. Introduction

This correspondence deals with the blind source separation (BSS) problem of a static system driven by (or linear mixtures of) independent source signals. To solve this problem, the ideas of the super-exponential methods (SEMs) in [1], [4], and [6] are used. Several researchers (e.g., [1], [4]–[6], [10]) have so far proposed some SEMs for solving independent component analysis (ICA), blind deconvolution (BD), and blind channel equalization (BCE). One of the attractive properties of the SEMs is that they are computationally efficient and that they converge to a desired solution at a super-exponential rate. However, almost all the conventional SEMs have the drawback that they are very sensitive to Gaussian noise (this will be shown in Section IV) because they utilize the second- and higher order cumulants of observations.

In this correspondence, we propose a new SEM that overcomes the drawback. The proposed SEM utilizes only the higher order cumulants of observations, and hence, the proposed SEM becomes robust to Gaussian noise, which is referred to as a *robust super-exponential method* (RSEM). Simulation results show that the proposed RSEM is robust to Gaussian noise and can successfully achieve the BSS of static systems (or linear mixtures of independent source signals).

Manuscript received December 2, 2003; revised June 9, 2004. This work was supported by the Scientific Research by the Ministry of Education, Culture, Sports, Science, and Technology of Japan under Grants-in-Aid 15700158 and 16500033. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Abdelhak M. Zoubir.

M. Kawamoto is with the Department of Electronic and Control Systems Engineering, Shimane University, Shimane 690-8504 Japan. He is also with Bio-Mimetic Control Research Center, RIKEN, Nagoya, 463-0003, Japan (e-mail: kawa@ecs.shimane-u.ac.jp).

K. Kohno is with the Department of Electronic and Control Systems Engineering, Shimane University, Shimane 690-8504 Japan (e-mail: kohno@yonago-k.ac.jp).

Y. Inouye is with the Department of Electronic and Control Systems Engineering, Shimane University, Shimane 690-8504 Japan. He is also with the Laboratory for Advanced Brain Signal Processing, Brain Science Institute, RIKEN, Wako-shi 351-0198, Japan (e-mail: inouye@riko.shimane-u.ac.jp).

Digital Object Identifier 10.1109/TSP.2005.845491

II. Problem Formulation

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

$$\boldsymbol{y}(t) = \boldsymbol{H}\boldsymbol{s}(t) + \boldsymbol{n}(t) \tag{1}$$

where $\boldsymbol{y}(t)$ represents an m-column output vector called the *observed signal*, $\boldsymbol{s}(t)$ represents an n-column input vector called the *source signal*, $\boldsymbol{H}$ is an $m \times n$ matrix, and $\boldsymbol{n}(t)$ 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:

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

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

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

where $\boldsymbol{g}_l = [g_{l1}, g_{l2}, \cdots, g_{ln}]^T := \boldsymbol{H}^T \boldsymbol{w}_l$ is an n-column vector. The BSS problem considered in this correspondence can be formulated as follows: Find n filters $\boldsymbol{w}_l$'s denoted by $\tilde{\boldsymbol{w}}_l$'s satisfying the following condition, without the knowledge of $\boldsymbol{H}$, even if the Gaussian noise $\boldsymbol{n}(t)$ is added to the observed signal $\boldsymbol{y}(t)$

$$\tilde{\boldsymbol{g}}_l = \boldsymbol{H}^T \tilde{\boldsymbol{w}}_l = \tilde{\boldsymbol{\delta}}_l, \quad l = 1, 2, \cdots, n \tag{4}$$

where $\tilde{\boldsymbol{\delta}}_l$ is an n-column vector whose elements $\tilde{\delta}_{lr}$ $(r = 1, 2, \cdots, n)$ are equal to zero, expect for the ρ_lth element, that is, $\tilde{\delta}_{lr} = c_l \delta(r - \rho_l)$, $r = 1, 2, \cdots, n$.

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 $\{1, 2, \cdots, n\}$ such that the set $\{\rho_1, \rho_2, \cdots, \rho_n\}$ is a permutation of the set $\{1, 2, \cdots, n\}$.

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

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

A2) The input sequence $\{\boldsymbol{s}(t)\}$ is a zero-mean, non-Gaussian vector stationary process whose element processes $\{s_i(t)\}$, $i = 1, 2, \cdots, n$ are mutually statistically independent and have nonzero $(p+1)$st-order cumulants κ_i defined as

$$\kappa_i = \operatorname{cum}\{\underbrace{s_i(t), s_i(t), \cdots, s_i(t)}_{p+1}\} \neq 0 \tag{5}$$

where $i = 1, 2, \cdots, n$, and $p \geq 2$.

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

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

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

III. Robust Super-Exponential Methods (RSEMs)

A. Two-Step Iterative Procedure for Vector $\boldsymbol{g}_l$

To find solutions in (4), the following two-step iterative procedure with respect to the elements $g_{lj}, j = 1, 2, \cdots, n$ of the vector $\boldsymbol{g}_l$ is used:

$$g_{lj}^{[1]} = \frac{\kappa_j}{\gamma_j a_j} g_{lj}^p, \quad j = 1, 2, \cdots, n \tag{6}$$

$$g_{lj}^{[2]} = \frac{g_{lj}^{[1]}}{\sqrt{\sigma_{z_l}^2}}, \quad j = 1, 2, \cdots, n \tag{7}$$

where g_{lj} on the right-hand side of (6) is an element of $\boldsymbol{g}_l$ before the iteration, $(\cdot)^{[1]}$ and $(\cdot)^{[2]}$ stand for the results of the first step and the second step per iteration, p is a non-negative integer, a_j denotes a positive number (in Section III-B, it will be shown how we choose the values of a_j's), γ_j denotes the fourth-order cumulant of $s_j(t)$, that is, γ_j is equal to κ_j in case of $p = 3$, and $\sigma_{z_l}^2$ denotes the variance of the output signal $z_l(t)$. Equation (6) is derived by replacing $\sigma_{s_j}^2$ of (26) in [1] with $\gamma_j a_j$, where $\sigma_{s_j}^2$ denotes the second-order cumulant of $s_i(t)$, and (7) is used to normalize $g_{lj}^{[1]}$ obtained by (6).

Here, it should be noted that in the conventional two-step procedures (e.g., [1], [4]–[6], [10]), the denominator of the right-hand side of (6) was set to 1 or the variance of $s_j(t)$, whereas we consider the fourth-order cumulant of $s_j(t)$, i.e., γ_j.

Let $g_{lj}(k)$ denote the value obtained in the kth cycle of the iterations of two steps (6) and (7). The important fact of the two-step procedure is that the n values $g_{lj}(k)$ $(j = 1, 2, \cdots, n)$ converge to zero, except for only one of the values, as the iteration number k approaches infinity, that is, $k \to \infty$. The magnitude of the remaining one converges to a positive constant. This will be shown in the following theorem.

Theorem 1: Let $g_{lj}(0)$ be an initial value for iterations of two steps (6) and (7) for each $j = 1, 2, \cdots, n$. Let α_j be a non-negative scalar defined as

$$\alpha_j = \left| \frac{\kappa_j}{\gamma_j a_j} \right|^{\frac{1}{p-1}}. \tag{8}$$

Let j_0 be $j_0 = \arg\max_{j \in \{1,2,\cdots,n\}} \alpha_j |g_{lj}(0)|$. Suppose the index j_0 is unique, that is, $\alpha_{j_0}|g_{lj_0}(0)| > \alpha_j |g_{lj}(0)|$ for any other $j \in \{1, 2, \cdots, n\}$. Then, as $k \to \infty$, it follows that

$$\lim_{k \to \infty} |g_{lj}(k)| = \begin{cases} 0, & \text{for } j \neq j_0 \\ \tilde{c}_j \neq 0, & \text{for } j = j_0 \end{cases} \tag{9}$$

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

Proof: From (6) and (8), choosing j_0 so that $g_{lj_0}(k) \neq 0$, we obtain

$$\frac{\left| g_{lj}^{[1]}(k) \right|}{\left| g_{lj_0}^{[1]}(k) \right|} = \frac{\alpha_j^{p-1}}{\alpha_{j_0}^{p-1}} \frac{\left| g_{lj}^{[1]}(k-1) \right|^p}{\left| g_{lj_0}^{[1]}(k-1) \right|^p} \tag{10}$$

where the integer k denotes the iteration time. Note that $|g_{lj}^{[1]}(k)|/|g_{lj_0}^{[1]}(k)|$ is not modified by the normalization of the second step. Therefore, it is possible to solve $|g_{lj}^{[2]}(k)|/|g_{lj_0}^{[2]}(k)|$ from the recursive formula (10), which yields

$$\frac{\left| g_{lj}^{[2]}(k) \right|}{\left| g_{lj_0}^{[2]}(k) \right|} = \frac{\alpha_{j_0}}{\alpha_j} \left(\frac{\alpha_j}{\alpha_{j_0}} \frac{\left| g_{lj}^{[2]}(0) \right|}{\left| g_{lj_0}^{[2]}(0) \right|} \right)^{p^k} \tag{11}$$

for any non-negative integer k. For $j_0 = \arg\max_j \alpha_j |g_{lj}(0)|$, one can see that all the other values $|g_{lj}^{[2]}(k)|$ $(j \neq j_0)$ quickly become small compared to $|g_{lj_0}^{[2]}(k)|$. Taking into account the normalization of the second step, this means that $|g_{lj_0}^{[2]}(k)| \neq 0$ and $|g_{lj}^{[2]}(k)| \to 0$ for all $j \neq j_0$. This implies that the infinite iteration of two steps (6) and (7) gives (9). Moreover, (11), along with the normalization of the second step, means that the sequence $\{g_{lj}(k)\}$ converges to a desired value at a super-exponential rate for all $j = 1, \cdots, n$. ■

B. Two-Step Iterative Procedure for Equalizer Vector $\boldsymbol{w}_l$

In (6) and (7), since the parameters g_{lj}'s include the unknown parameters h_{ij}'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 $\boldsymbol{w}_l$ so that the two steps (6) and (7) can be handled indirectly.

$$\min_{\boldsymbol{w}_l} (\boldsymbol{H}^T \boldsymbol{w}_l - \boldsymbol{g}_l)^T \boldsymbol{\Lambda} (\boldsymbol{H}^T \boldsymbol{w}_l - \boldsymbol{g}_l), \quad l = 1, 2, \cdots, n. \tag{12}$$

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

$$\boldsymbol{w}_l = (\boldsymbol{H}\boldsymbol{\Lambda}\boldsymbol{H}^T)^\dagger \boldsymbol{H}\boldsymbol{\Lambda}\boldsymbol{g}_l, \quad l = 1, 2, \cdots, n \tag{13}$$

where $\dagger$ denotes the pseudo-inverse operation of a matrix. In the conventional methods [1], [4]–[6], [10], the positive diagonal elements of $\boldsymbol{\Lambda}$ are set to 1 or the variances of the source signals. This means that $\boldsymbol{H}\boldsymbol{\Lambda}\boldsymbol{H}^T$ is calculated by the second-order statistics of the observed signal $\boldsymbol{y}(t)$. We consider that this is the reason why the conventional methods are sensitive to Gaussian noise.

In what follows, we will show that the weighted least squares approach in (12) can be applied to a set of fourth-order cumulants of the observed signals $y_i(t)$ $(i = 1, 2, \cdots, m)$, if we choose appropriately a diagonal matrix $\boldsymbol{\Lambda}$ in (12). To this end, we introduce fourth-order cumulant matrices of the m-vector random process $\{\boldsymbol{y}(t)\}$ [8], which constitute a set of $m \times m$ matrices $\boldsymbol{C}_{\boldsymbol{y},i,j}^{(4)}$ $(i, j = 1, 2, \cdots, m)$, where each is defined by

$$[\boldsymbol{C}_{\boldsymbol{y},i,j}^{(4)}]_{q,r} = \text{cum}\{y_q(t), y_r(t), y_i(t), y_j(t)\} \tag{14}$$

where $[x]_{q,r}$ denotes the (q, r)th element of the matrix $\boldsymbol{C}_{\boldsymbol{y},i,j}^{(4)}$. Then, we consider an $m \times m$ matrix $\tilde{\boldsymbol{R}}$ expressed by

$$\tilde{\boldsymbol{R}} = \sum_{i,j=1}^{m} \beta_{ij} \boldsymbol{C}_{\boldsymbol{y},i,j}^{(4)} \tag{15}$$

where β_{ij}'s are either 1 or 0, which represent *design parameters*. It is shown by a simple calculation (see the Appendix) that (15) becomes

$$\tilde{\boldsymbol{R}} = \boldsymbol{H}\tilde{\boldsymbol{\Lambda}}\boldsymbol{H}^T \tag{16}$$

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

$$\tilde{\boldsymbol{\Lambda}} := \text{diag}\{\gamma_1 \tilde{a}_1, \gamma_2 \tilde{a}_2, \cdots, \gamma_n \tilde{a}_n\} \tag{17}$$

$$\tilde{a}_r := \sum_{i,j=1}^{m} \beta_{ij} h_{ir} h_{jr}, \quad r = 1, 2, \cdots, n \tag{18}$$

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

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

$$\hat{\boldsymbol{\Lambda}} := \text{diag}\,\{|\gamma_1\tilde{a}_1|, |\gamma_2\tilde{a}_2|, \cdots, |\gamma_n\tilde{a}_n|\} \tag{19}$$

is positive semi-definite. It is clear from the definitions (17) and (19) that there exists a sign matrix $\dot{\boldsymbol{I}}$ such that $\tilde{\boldsymbol{\Lambda}} := \hat{\boldsymbol{\Lambda}}\dot{\boldsymbol{I}}$, where the sign matrix $\dot{\boldsymbol{I}}$ is defined as a diagonal matrix whose diagonal elements are either $+1$ or -1.

Remark 1: If $\hat{\boldsymbol{\Lambda}}$ is full rank, then by putting $\boldsymbol{\Lambda}$ in (12) to $\hat{\boldsymbol{\Lambda}}$, the solution $(\boldsymbol{H}\hat{\boldsymbol{\Lambda}}\boldsymbol{H}^T)^\dagger\boldsymbol{H}\hat{\boldsymbol{\Lambda}}\boldsymbol{g}_l$ can be obtained from (12). However, $\boldsymbol{H}\hat{\boldsymbol{\Lambda}}\boldsymbol{H}^T$ cannot be calculated from (15), that is, $\tilde{\boldsymbol{R}} \neq \boldsymbol{H}\hat{\boldsymbol{\Lambda}}\boldsymbol{H}^T$.

Here, we show the following theorem.

Theorem 2: If $\boldsymbol{H}$ is of full column rank and both $\hat{\boldsymbol{\Lambda}}$ and $\tilde{\boldsymbol{\Lambda}}$ are of full rank, then

$$(\boldsymbol{H}\hat{\boldsymbol{\Lambda}}\boldsymbol{H}^T)^\dagger\boldsymbol{H}\hat{\boldsymbol{\Lambda}}\boldsymbol{g}_l = (\boldsymbol{H}\tilde{\boldsymbol{\Lambda}}\boldsymbol{H}^T)^\dagger\boldsymbol{H}\tilde{\boldsymbol{\Lambda}}\boldsymbol{g}_l \tag{20}$$

where $\tilde{\boldsymbol{\Lambda}} := \hat{\boldsymbol{\Lambda}}\dot{\boldsymbol{I}}$.

Proof: Let the left-hand and the right-hand sides of (20) be denoted by $\hat{\boldsymbol{w}}_l$ and $\tilde{\boldsymbol{w}}_l$, respectively. Then, we will show that $\tilde{\boldsymbol{w}}_l$ can be derived from $\hat{\boldsymbol{w}}_l = (\boldsymbol{H}\hat{\boldsymbol{\Lambda}}\boldsymbol{H}^T)^\dagger\boldsymbol{H}\hat{\boldsymbol{\Lambda}}\boldsymbol{g}_l$.

Since $\boldsymbol{H}$ has full column rank, using a property of the pseudo-inverse operation (see [3, p. 433]), we obtain

$$\begin{aligned}\hat{\boldsymbol{w}}_l &:= (\boldsymbol{H}\hat{\boldsymbol{\Lambda}}\boldsymbol{H}^T)^\dagger\boldsymbol{H}\hat{\boldsymbol{\Lambda}}\boldsymbol{g}_l = \boldsymbol{H}^{T\dagger}(\boldsymbol{H}\hat{\boldsymbol{\Lambda}})^\dagger\boldsymbol{H}\hat{\boldsymbol{\Lambda}}\boldsymbol{g}_l \\ &= \boldsymbol{H}^{T\dagger}\hat{\boldsymbol{\Lambda}}^{-1}\boldsymbol{H}^\dagger\boldsymbol{H}\hat{\boldsymbol{\Lambda}}\boldsymbol{g}_l = \boldsymbol{H}^{T\dagger}\boldsymbol{g}_l\end{aligned} \tag{21}$$

where the fourth equality comes from the fact that $\boldsymbol{H}^\dagger\boldsymbol{H} = \boldsymbol{I}$ because $\boldsymbol{H}$ is of full column rank. From (21) and $\boldsymbol{H}^\dagger\boldsymbol{H} = \boldsymbol{I}$, we have

$$\begin{aligned}\boldsymbol{H}^{T\dagger}\boldsymbol{g}_l &= \boldsymbol{H}^{T\dagger}\tilde{\boldsymbol{\Lambda}}^{-1}\boldsymbol{H}^\dagger\boldsymbol{H}\tilde{\boldsymbol{\Lambda}}\boldsymbol{g}_l \\ &= (\boldsymbol{H}\tilde{\boldsymbol{\Lambda}}\boldsymbol{H}^T)^\dagger\boldsymbol{H}\tilde{\boldsymbol{\Lambda}}\boldsymbol{g}_l = \tilde{\boldsymbol{w}}_l\end{aligned} \tag{22}$$

where the last equality comes from the definition of $\tilde{\boldsymbol{w}}_l$. The reverse, which $\hat{\boldsymbol{w}}_l$ can be derived from $\tilde{\boldsymbol{w}}_l = (\boldsymbol{H}\tilde{\boldsymbol{\Lambda}}\boldsymbol{H}^T)^\dagger\boldsymbol{H}\tilde{\boldsymbol{\Lambda}}\boldsymbol{g}_l$, can also be shown in the same way. Therefore, both $\hat{\boldsymbol{w}}_l$ and $\tilde{\boldsymbol{w}}_l$ are identical. ■

Remark 2: If $\boldsymbol{H}$ is not of full column rank, Theorem 2 does not hold. Because, in such a case, $\boldsymbol{H}^T\boldsymbol{H}$ does not become a nonsingular matrix. Moreover, it can be seen from (21) and (22) that $\hat{\boldsymbol{w}}_l$ and $\tilde{\boldsymbol{w}}_l$ are, respectively, irrelevant to $\hat{\boldsymbol{\Lambda}}$ and $\tilde{\boldsymbol{\Lambda}}$, that is, Theorem 2 holds for any pair of full-rank diagonal matrices. In fact, $\hat{\boldsymbol{w}}_l = \tilde{\boldsymbol{w}}_l = \boldsymbol{H}^{T\dagger}\boldsymbol{g}_l$, which is shown in (21) and (22), attains the zero minimum value of the weighted least squares function in (12) for any diagonal positive definite matrix. In general, the right-hand side of (20) is always expressed by the fourth-order cumulants or fourth- and higher order cumulants of $\{\boldsymbol{y}(t)\}$.

From Theorem 2 and Remark 2, it is seen that the right-hand side of (13) can be given by the right-hand side of (20) under the condition that the diagonal matrix $\tilde{\boldsymbol{\Lambda}}$ $(= \hat{\boldsymbol{\Lambda}}\dot{\boldsymbol{I}})$ is full rank. This condition, however, will be satisfied by the following theorem.

Theorem 3: Let $\boldsymbol{H}$ be full column rank and γ_i $(i = 1, 2, \cdots, n)$ be nonzero for all i. Suppose that $\beta_{ij} = 1$ for $i = j$ and $\beta_{ij} = 0$ for $i \neq j$ [see (15)]. Then, the diagonal matrix $\tilde{\boldsymbol{\Lambda}}$ in (17) becomes full rank.

Proof: If β_{ij} in (15) is 1 for $i = j$ and 0 for $i \neq j$, then $\tilde{a}_r$'s of $\tilde{\boldsymbol{\Lambda}}$ in (17) become

$$\tilde{a}_r = \sum_{i=1}^{m} h_{ir}^2, \quad \text{for } r = 1, 2, \cdots, n. \tag{23}$$

Suppose $\tilde{\boldsymbol{\Lambda}}$ does not have full rank. Then, one of the diagonal elements of $\tilde{\boldsymbol{\Lambda}}$ becomes zero, that is, $\gamma_r\tilde{a}_r = 0$ for some r. It implies that $\tilde{a}_r = \sum_{i=1}^{m} h_{ir}^2 = 0$ because $\gamma_r \neq 0$. If $\sum_{i=1}^{m} h_{ir}^2 = 0$, then $h_{ir} = 0$ for all i. This contradicts the assumption that $\boldsymbol{H}$ is full column rank. Therefore, $\tilde{\boldsymbol{\Lambda}}$ under these conditions is of full rank. ■

For the time being, in this correspondence, we consider (15) with $\beta_{ij} = 1$ for $i = j$ and $\beta_{ij} = 0$ for $i \neq j$. As for $\boldsymbol{H}\tilde{\boldsymbol{\Lambda}}\boldsymbol{g}_l$, by using (6) with $a_j = \tilde{a}_j$ in (23) and the similar way as in [1], it can be calculated by

$$\boldsymbol{d}_l := [d_{l1}, d_{l2}, \cdots, d_{lm}]^T \tag{24}$$

where d_{lj} is given by $d_{lj} = \text{cum}\{\underbrace{z_l(t), z_l(t), \cdots, z_l(t)}_{p}, y_j(t)\}$. Then, (13) can be expressed as

$$\boldsymbol{w}_l^{[1]} := \tilde{\boldsymbol{R}}^\dagger\boldsymbol{d}_l, \quad l = 1, 2, \cdots, n. \tag{25}$$

Since the second step (7) is a normalization of $\boldsymbol{g}_l$, it is easily shown that the second step reduces to

$$\boldsymbol{w}_l^{[2]} := \frac{\boldsymbol{w}_l^{[1]}}{\sqrt{\sigma_{z_l}^2}}. \quad l = 1, 2, \cdots, n \tag{26}$$

Therefore, (25) and (26) are our proposed two steps to modify $\boldsymbol{w}_l$, which becomes one cycle of iterations in the super-exponential method [1], [4]–[6], [10]. Then, since the right-hand side of (25) consists of only higher order cumulants, the modification of $\boldsymbol{w}_l$ is not affected by Gaussian noise. This comes from the fact that higher order cumulants are insensitive to additive (even colored) Gaussian noise (see [8, Prop. 4, p. 2463]). This is a *novel key point* of our proposed super-exponential method, from which the proposed method is referred to as a robust super-exponential method (RSEM).

C. Proposed RSEM

For now, there are two approaches to multichannel (or MIMO) BSS, a *concurrent BSS approach* and a *deflationary BSS approach*. The former is to find all the n filters $\tilde{\boldsymbol{w}}_l$'s in (4) concurrently, whereas the latter finds sequentially (or iteratively with respect to source signals) the filters $\tilde{\boldsymbol{w}}_l$'s one by one. It is well known that iterative algorithms based on the former approach converge to a desired solution when they start in a neighborhood of the desired solution, whereas iterative algorithms based on the latter approach converge to a desired solution globally (or regardless of their initialization) [1]. The latter approach is employed in this correspondence. Let l denote the number of the sources to be extracted. At first, set $l = 1$; then, $\tilde{\boldsymbol{w}}_1$ is calculated by the two steps (25) and (26) such that $\boldsymbol{H}^T\tilde{\boldsymbol{w}}_1 = \tilde{\boldsymbol{\delta}}_1 = [0, \cdots, 0, 1(\rho_1\text{th element}), 0, \cdots, 0]^T$. Next, the contribution signals $v_{i\rho_1}(t) = h_{i\rho_1}s_{\rho_1}(t)$ $(i = 1, 2, \cdots, m)$ are calculated by using the output signal $z_1(t) = \tilde{\boldsymbol{w}}_1^T\boldsymbol{y}(t)$. Then, by calculating $y_i(t) - v_{i\rho_1}(t)$ for $i = 1, 2\cdots, m$, we remove the contribution signals from the outputs in order to define the outputs of a multichannel system with $n - 1$ inputs and m outputs. The number of inputs becomes deflated by one. The procedures mentioned above are continued until $l = n$. Therefore, the proposed RSEM is summarized as shown in Table I.

The procedure from Steps 5 to 7 are implemented to make it possible to obtain solutions in (4). In Step 6, the calculation of $\boldsymbol{y}_l(t) - (\boldsymbol{d}_l/\kappa_{\tilde{z}})\tilde{z}_l(t)$ is equivalent to the calculations of $y_i(t) - v_{i\rho_1}(t)$ $(i = 1, 2, \cdots, n)$ mentioned above. (On the details of Step 6, see Section V or [5].)

TABLE I
PROPOSED METHOD

Step	Contents
1	Set $l = 1$ (where l denotes the number of the source signals separated).
2	Choose random initial value $\boldsymbol{w}_l(0)$, where $\boldsymbol{w}_l(0)$ denotes the initial value of $\boldsymbol{w}_l$. Set k in $\boldsymbol{w}_l(k)$ to 0 (k denotes the iteration number).
3	Calculate (2).
4	Calculate $\boldsymbol{w}_l(k)$ using (25) and (26).
5	After $k \to \infty$, $\boldsymbol{d}_l$ is calculated by using $\tilde{z}_l(t)$ which is estimated by using $\boldsymbol{w}_l(\infty)$.
6	Calculate $\tilde{\boldsymbol{y}}_l(t) = \boldsymbol{y}_l(t) - (\boldsymbol{d}_l/\kappa_{\tilde{z}})\tilde{z}_l(t)$, where $\boldsymbol{y}_l(t)$ is the observed signal in number l and $\kappa_{\tilde{z}_l}$ is the $(p+1)$st-order cumulant of $\tilde{z}_l(t)$.
7	If the subscript l of $\boldsymbol{y}_l(t)$ is less then n, then set $\boldsymbol{y}(t) = \tilde{\boldsymbol{y}}_l(t)$, $l = l + 1$, and the procedures (Step 2 to Step 6) are continued until $l = n$.

IV. SIMULATION RESULTS

To demonstrate the validity of the proposed RSEM, many computer simulations were conducted. One of the results is shown in this section. We considered a two-input and three-output system, that is, $\boldsymbol{H}$ in (1) was set to be

$$\boldsymbol{H} = \begin{bmatrix} 1.0 & 0.6 \\ 0.7 & 1.0 \\ 0.2 & 0.5 \end{bmatrix}. \tag{27}$$

Two source signals $s_1(t)$ and $s_2(t)$ were sub-Gaussian and super-Gaussian, respectively, in which $s_1(t)$ takes one of two values -1 and 1 with equal probability 1/2, $s_2(t)$ takes one of three values $-2, 0$, and 2 with probability 1/8, 6/8, and 1/8, respectively, and they are zero-mean and unit variance. The parameter p in (5) was set to be $p = 3$, that is, κ_j $(j = 1, 2)$ in (6), were the fourth-order cumulants of the source signals. These values were set to be $\kappa_1 = -2$ and $\kappa_2 = 1$. 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. Initial values of $\boldsymbol{w}_l$ were randomly chosen from the values between -1 and 1. As a measure of performance, we used the *multichannel intersymbol interference* ($\mathrm{M_{ISI}}$) defined in the logarithmic (decibel) scale by

$$\mathrm{M_{ISI}} = 10\log_{10}\left[\sum_{l=1}^{n}\frac{\left|\sum_{j=1}^{n}|g_{lj}|^2 - |g_{l\cdot}|_{\max}^2\right|}{|g_{l\cdot}|_{\max}^2} + \sum_{j=1}^{n}\frac{\left|\sum_{l=1}^{n}|g_{lj}|^2 - |g_{\cdot j}|_{\max}^2\right|}{|g_{\cdot j}|_{\max}^2}\right] \tag{28}$$

where $|g_{l\cdot}|_{\max}^2$ and $|g_{\cdot j}|_{\max}^2$ are, respectively, defined by $|g_{l\cdot}|_{\max}^2 := \max_{j=1,\cdots,n}|g_{lj}|^2$ and $|g_{\cdot j}|_{\max}^2 := \max_{l=1,\cdots,n}|g_{lj}|^2$. The value of $\mathrm{M_{ISI}}$ becomes $-\infty$, if the $\tilde{\boldsymbol{g}}_l$'s in (4) are obtained, and hence, a minus large value of $\mathrm{M_{ISI}}$ indicates the proximity to the desired solution. As a conventional method, the method proposed in [5] was used for comparison.

Fig. 1 shows the results of performances for the proposed RSEM and the conventional SEM when the SNR levels were, respectively, taken to be 0 $(\sigma_n^2 = 1)$, 2.5, 5, 10, 15, and ∞ dB $(\sigma_n^2 = 0)$, in which each $\mathrm{M_{ISI}}$ shown in Fig. 1 was the average of the performance results obtained by 50 independent Monte Carlo runs. In each Monte Carlo run, the number of the integers k's in Step 4 (see Table I) was 10, in which $\hat{\boldsymbol{R}}$ and $\boldsymbol{d}_l$ were estimated by data samples in the following three cases: (Case 1) 1000 data, (Case 2) 10 000 data, and (Case 3) 100 000 data.

It can be seen from Fig. 1 that as the number of data samples that are needed to estimate the cumulants increases, the proposed RSEM shows better performance, whereas the performances of the conventional SEM hardly change. This implies that the performance of the RSEM depends on the accuracy of the estimate of the higher order cumulants. We consider, however, that since, in the above three cases, the performances of the RSEM are better than the ones of the conventional SEM, the proposed RSEM is effective for solving the BSS problem.

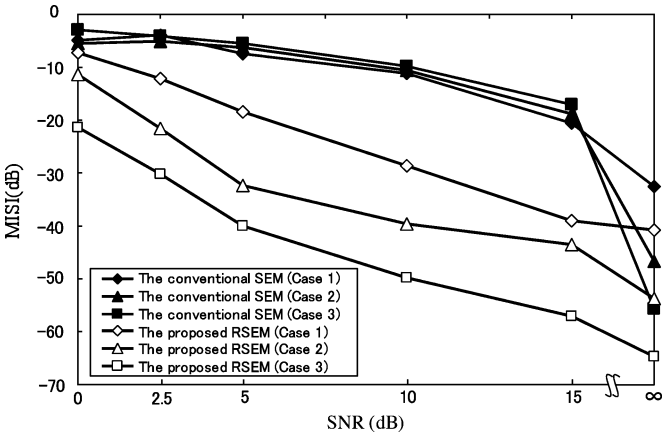

Fig. 1. Performances for the proposed RSEM and the conventional SEM.

V. DISCUSSIONS

In the method shown in Table I, the calculation of Step 6 is important for implementing a *deflationary* SEM. Let us review the calculation of Step 6 in Table I. Suppose that $\tilde{\boldsymbol{g}}_1 = \tilde{\boldsymbol{\delta}}_1$ in (4) is obtained for $l = 1$. Then, from (3)–(5), and (24), $\tilde{z}_1(t)$, $\kappa_{\tilde{z}_1}$, and $\boldsymbol{d}_1$ are, respectively

$$\tilde{z}_1(t) = c_1 s_{\rho_1}(t) + \tilde{\boldsymbol{w}}_1^T \boldsymbol{n}(t) \tag{29}$$

$$\kappa_{\tilde{z}_1} = c_1^{p+1}\kappa_{\rho_1} \tag{30}$$

$$\boldsymbol{d}_1 = [h_{1\rho_1}c_1^p\kappa_{\rho_1}, h_{2\rho_1}c_1^p\kappa_{\rho_1}, \cdots, h_{m\rho_1}c_1^p\kappa_{\rho_1}]^T. \tag{31}$$

Therefore, from (29)–(31), $\hat{\boldsymbol{y}}_1(t) := (\boldsymbol{d}_1/\kappa_{\tilde{z}})\tilde{z}_1(t)$ of Step 6 in Table I becomes

$$\hat{\boldsymbol{y}}_1(t) = \begin{bmatrix} h_{1\rho_1}s_{\rho_1}(t) + \eta_{11}\tilde{\boldsymbol{w}}_1^T\boldsymbol{n}(t) \\ h_{2\rho_1}s_{\rho_1}(t) + \eta_{21}\tilde{\boldsymbol{w}}_1^T\boldsymbol{n}(t) \\ \vdots \\ h_{m\rho_1}s_{\rho_1}(t) + \eta_{m1}\tilde{\boldsymbol{w}}_1^T\boldsymbol{n}(t) \end{bmatrix} \tag{32}$$

where $\eta_{i1} := h_{i\rho_1}/c_1$ $(i = 1, 2, \cdots, m)$. When we calculate

$$\tilde{\boldsymbol{y}}_1(t) = \boldsymbol{y}_1(t) - \hat{\boldsymbol{y}}_1(t) \tag{33}$$

the output $\tilde{\boldsymbol{y}}_1(t)$ in (33) is the output of a system that has $n-1$ inputs and m outputs. Therefore, by the calculation of Step 6, the number of inputs becomes deflated by one.

VI. Conclusions

We have proposed a deflationary SEM for solving the BSS problem, in which the solutions of the problem $\tilde{\boldsymbol{w}}_l$'s satisfying (4) are found one by one. The proposed SEM is not sensitive to Gaussian noise, which is referred to as a robust super-exponential method (RSEM). This is a novel property of the proposed method, whereas the conventional methods do not posses it. It was shown from the simulation results that the proposed RSEM was robust to Gaussian noise and could successfully solve the BSS problem.

Appendix
Derivation of (16)

From the properties of the cumulant (see [6]), $\text{cum}\{y_q(t), y_r(t), y_i(t), y_j(t)\}$ in (14) becomes

$$\begin{aligned}&\text{cum}\{y_q(t), y_r(t), y_i(t), y_j(t)\}\\&= \sum_{l_1,l_2,l_3,l_4} h_{ql_1}h_{rl_2}h_{il_3}h_{jl_4}\text{cum}\{s_{l_1}(t), s_{l_2}(t), s_{l_3}(t), s_{l_4}(t)\}\\&\quad + \text{cum}\{n_q(t), n_r(t), n_i(t), n_j(t)\} \qquad (34)\\&= \sum_{l=1}^{n} h_{ql}h_{rl}h_{il}h_{jl}\gamma_l = \boldsymbol{h}_q^T\boldsymbol{\Lambda}_{\gamma h_{ij}}\boldsymbol{h}_r \qquad (35)\end{aligned}$$

where the second equality comes from assumptions A2) and A3) and the fact that the fourth-order cumulant of Gaussian noises $n_i(t)$'s are equal to zero, $\boldsymbol{h}_q := [h_{q1}, h_{q2}, \cdots, h_{qn}]^T$, $\boldsymbol{\Lambda}_{\gamma h_{ij}}$ is a diagonal matrix defined by

$$\boldsymbol{\Lambda}_{\gamma h_{ij}} = \begin{bmatrix} \gamma_1 h_{i1}h_{j1} & 0 & \cdots & 0 \\ 0 & \gamma_2 h_{i2}h_{j2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \gamma_n h_{in}h_{jn} \end{bmatrix}$$

and $\boldsymbol{h}_r := [h_{r1}, h_{r2}, \cdots, h_{rn}]^T$. From (35), we obtain

$$\sum_{i,j=1}^{m} \beta_{ij}\text{cum}\{y_q(t)y_r(t)y_i(t)y_j(t)\} = \boldsymbol{h}_q^T\boldsymbol{\Lambda}_{\gamma\Sigma h_{ij}}\boldsymbol{h}_r \qquad (36)$$

where $\boldsymbol{\Lambda}_{\gamma\Sigma h_{ij}}$ is a diagonal matrix defined by

$$\begin{bmatrix} \gamma_1\sum_{i,j} h_{i1}h_{j1} & 0 & \cdots & 0 \\ 0 & \gamma_2\sum_{i,j} h_{i2}h_{j2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \gamma_n\sum_{i,j} h_{in}h_{jn} \end{bmatrix}.$$

It can be seen that (36) expresses the (q,r)th element of $\boldsymbol{H}\tilde{\boldsymbol{\Lambda}}\boldsymbol{H}^T$. Therefore, (16) holds true.

Acknowledgment

The authors would like to thank Dr. M. Ohata, RIKEN, and M. Ito, Nagoya University for giving useful suggestions.

References

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

[2] M. Kawamoto and Y. Inouye, "A deflation algorithm for the blind source-factor separation of MIMO-FIR systems driven by colored sources," *IEEE Signal Process. Lett.*, vol. 10, no. 11, pp. 343–346, Nov. 2003.

[3] P. Lancaster and M. Tismenetsky, *The Theory of Matrices*, Second ed. New York: Academic, 1985.

[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] ——, "Fast adaptive super-exponential multistage beamforming cellular base-station transceivers with antenna arrays," *IEEE Trans. Veh. Technol.*, vol. 48, no. 4, pp. 1017–1028, Jul. 1999.

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

[7] C. Simon, P. Loubaton, and C. Jutten, "Separation of a class of convolutive mixtures: a contrast function approach," *Signal Process.*, vol. 81, pp. 883–887, 2001.

[8] L. Tong, Y. Inouye, and R.-w. Liu, "Waveform-preserving blind estimation of multiple independent sources," *IEEE Trans. Signal Process.*, vol. 41, no. 7, pp. 2461–2470, Jul. 1993.

[9] J. K. Tugnait, "Identification and deconvolution of multichannel non-Gaussian processes using higher order statistics and inverse filter criteria," *IEEE Trans. Signal Process.*, vol. 45, no. 3, pp. 658–672, Mar. 1997.

[10] K. L. Yeung and S. F. Yau, "A cumulant-based super-exponential algorithm for blind deconvolution of multi-input multi-output systems," *Signal Process.*, vol. 67, pp. 141–162, 1998.

Iterative Decoding of Wrapped Space-Time Codes

Aydin Sezgin, *Student Member, IEEE,* and Holger Boche

***Abstract*—We study the iterative decoding of Wrapped Space-Time Codes (WSTCs) employing per-survivor-processing with the soft-output Viterbi-algorithm (SOVA). We use a novel receiver scheme that incorporates extrinsic information delivered by the SOVA. The decision metric of the SOVA is developed, and the performance is analyzed.**

***Index Terms*—Iterative decoding, MIMO, SOVA, space-time codes.**

I. Introduction

In recent years, the goal of providing high-speed wireless data services has generated a great amount of interest among the research community. Recent information-theoretic results have demonstrated that the capacity of the system in the presence of Rayleigh fading improves significantly with the use of multiple transmit and receive antennas [1], [2].

Diagonal Bell Labs Layered Space-Time (DBLAST), which is an architecture that theoretically achieves a capacity for such multiple-input

Manuscript received January 22, 2004; revised June 8, 2004. This work was supported in part by the German ministry of education and research (BMBF) under Grant 01BU150. Part of this work was presented at the IEEE International Symposium on Personal, Indoor, and Mobile Radio Communications, Beijing, China, September 7–10, 2003. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Vikram Krishnamurthy.

A. Sezgin is with the Department of Broadband Mobile Communication Networks, Fraunhofer Institute for Telecommunications, HHI, D-10587 Berlin, Germany (e-mail: aydin.sezgin@hhi.fhg.de).

H. Boche is with the Department of Mobile Communication Networks, Technical University of Berlin, D-10587 Berlin, Germany (e-mail: holger.boche@mk.tu-berlin.de).

Digital Object Identifier 10.1109/TSP.2005.845493