Robust Super-Exponential Methods for Blind Equalization in the Presence of Gaussian Noise

Mitsuru Kawamoto, *Member, IEEE*, Masashi Ohata, *Member, IEEE*, Kiyotaka Kohno, *Student Member, IEEE*, Yujiro Inouye, *Member, IEEE*, and Asoke K. Nandi, *Senior Member, IEEE*

***Abstract*—This paper deals with the blind equalization problem of a single-input single-output infinite-impulse-response (SISO-IIR) system with additive Gaussian noise. To solve the problem, we propose a "super-exponential method" (SEM). The novel point of the proposed SEM is that, even when Gaussian noise is added to the output of the system, the blind equalization can be achieved with as little influence of Gaussian noise as possible; hence the proposed SEM is referred to as a "robust super-exponential method" (RSEM). Simulation results show the validity of the proposed RSEM.**

***Index Terms*—Blind equalization, Gaussian noise, single-input single-output infinite-impulse-response (SISO-IIR) systems, super-exponential methods (SEMs).**

I. Introduction

In applications such as mobile or wireless communications, an input signal often propagates through a multipath environment of an unknown transfer function between the signal source and a receiver. Blind equalization is used to reconstruct the original input signal and/or to estimate the transfer function from the received signal [1].

Recently, Shalvi and Weinstein proposed an attractive approach for the blind equalization of single-input single-output (SISO) systems, which is called the super-exponential method (SEM) [9], and then several researchers extended the idea of the SEM, e.g., see [5], [7], [11], and [12] and references therein. One of the attractive properties of the SEM is to converge iteratively at a super-exponential rate to a desired solution which achieves the blind equalization; hence the "super-exponential" method was named. However, the SEMs have such a significant drawback that, if the SEMs are applied to the blind equalization in the presence of additive Gaussian noise, then the convergence of the SEMs close to the desired solutions cannot be guaranteed [9]. Such algorithms based on the second-order statistics as the fixed-point Bussgang [2] are also sensitive to Gaussian noise.

Manuscript received October 18, 2004; revised February 18, 2005. This paper was recommended by Associate Editor B. C. Levy.

M. Kawamoto is with the Department of Electrical Engineering and Electronics, University of Liverpool, L69 3GJ Liverpool, U.K., on sabbatical from the Department of Electronic and Control Systems Engineering, Shimane University, Shimane 690-8504, Japan (e-mail: kawa@ecs.shimane-u.ac.jp).

M. Ohata is with Bio-Mimetic Control Research Center, RIKEN, Nagoya 463 003, Japan (ohatama@bmc.riken.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 351-098, Japan (e-mail: inouye@riko.shimane-u.ac.jp).

A. K. Nandi is with the Department of Electrical Engineering and Electronics, University of Liverpool, L69 3GJ Liverpool, U.K. (a.nandi@liverpool.ac.uk).

Digital Object Identifier 10.1109/TCSII.2005.852174

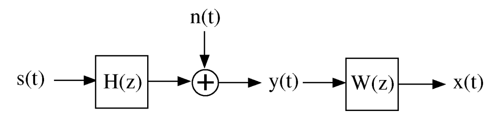

Fig. 1. Composite system of an unknown system and a filter.

In this paper, an approach is proposed in order to overcome the drawback of the SEMs. In the proposed approach, only higher order cumulants are used; consequently, the proposed algorithm can be used to detect the desired solutions with as little influence of Gaussian noise as possible, from which the proposed SEM is referred to as a *robust super-exponential method* (RSEM). Computer simulations are presented to demonstrate the validity of the proposed RSEM.

II. Problem Formulation and Assumptions

We consider a SISO system with an additive noise as described by

$$y(t) = \sum_{k=-\infty}^{\infty} h^{(k)} s(t-k) + n(t) \tag{1}$$

where $\{s(t)\}$ is an unobserved input sequence generated from a discrete-time stationary random process, $h^{(k)}$ is the impulse response of an unknown time-invariant system defined by $H(z) = \sum_{k=-\infty}^{\infty} h^{(k)} z^k$, and $y(t)$ and $n(t)$ denote the output of the system and Gaussian noise, respectively. Fig. 1 illustrates a diagram of the basic problem. Namely, our objective in this paper is to propose a method for adjusting the equalizer $W(z) = \sum_{k=L_1}^{L_2} w^{(k)} z^k$ so that $G(z) := W(z)H(z)$ becomes

$$\hat{G}(z) = \hat{W}(z)H(z) = cz^{k_1} \tag{2}$$

even if the Gaussian noise $n(t)$ is included into the output $y(t)$, where c in cz^{k_1} is a nonzero complex number standing for a scale change and a phase shift, and the superscript "k_1" of z^{k_1} denotes an integer standing for a constant delay. Note that the notation z is used instead of the commonly used z^{-1} in the z-transform [4]. We allow all of the above signals and the parameters of the system and equalizer to be complex-valued.

To find the solution (2), we put the following assumptions on the system, the input signal, and the equalizer.

A1) The unknown system $H(z)$ is a stable, possibly nonminimum-phase, linear time-invariant filter whose inverse (which may be noncausal and stable) $H(z)^{-1}$ exists.

A2) The input sequence $\{s(t)\}$ is a complex, zero-mean, non-Gaussian random process. Moreover, the process $\{s(t)\}$ is an i.i.d. process with a nonzero variance σ_s^2

and a nonzero $(p+q+1)$st-order cumulants, κ_s defined as

$$\kappa_s = \text{cum}\left\{\underbrace{s(t),\cdots,s(t)}_{p},\underbrace{s^*(t),\cdots,s^*(t)}_{q_0+1}\right\} \tag{3}$$

where p and q are nonnegative integers such that $(p+q) \geq 2$, and $\text{cum}\{s_1, s_2, \cdots, s_n\}$ denotes the nth-order (joint) cumulant of $s_1, s_2, \cdots, s_n$.

A3) The equalizer $W(z) = \sum_{k=L_1}^{L_2} w^{(k)} z^k$ is a finitie-impulse-response (FIR) system of sufficient length $L = L_2 - L_1 + 1$ so that the truncation effect can be ignored.

The combined system response subject to the finite length restriction is

$$g^{(k)} = \sum_{l=L_1}^{L_2} w^{(l)} h^{(k-l)}. \tag{4}$$

In a vector notation, (4) can be rewritten as

$$\boldsymbol{g} = \boldsymbol{H}\boldsymbol{w} \tag{5}$$

where $\boldsymbol{g}$ is a possibly infinite vector of the combined system $\boldsymbol{g} = [\cdots, g^{(-1)}, g^{(0)}, g^{(1)}, \cdots]^T$, $\boldsymbol{w}$ is an L-column vector, that is, $\boldsymbol{w} = [w^{(L_1)}, w^{(L_1+1)}, \cdots, w^{(L_2)}]^T$, and $\boldsymbol{H} = [h_{kl}]$ is a matrix of L columns and possibly infinite number of rows, whose elements are $h_{kl} = h^{(k-l)}$, $k = -\infty, \cdots, \infty$, $l = L_1, (L_1+1), \cdots, L_2$.

III. Robust Super-Exponential Methods

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

To find the solution in (2), the following two-step iterative procedure with respect to the elements $g^{(k)}$'s of $\boldsymbol{g}$ is used:

$$g^{(k)[1]} = \frac{\kappa_s}{\gamma_s \alpha_k}\left(g^{(k)}\right)^p \left(g^{(k)*}\right)^q \tag{6}$$

$$g^{(k)[2]} = \frac{g^{(k)[1]}}{\sqrt{\sigma_x^2}} \tag{7}$$

where $(\cdot)^{[1]}$ and $(\cdot)^{[2]}$ stand for the results of the first step and the second step per iteration, $g^{(k)}$ in the right-hand side of (6) is $g^{(k)[2]}$ at the previous step [note that, at first iteration, $g^{(k)}$ in the right-hand side of (6) is an initial value of $g^{(k)}$], p and q are nonnegative integers such that $(p+q) \geq 2$, γ_s denotes the fourth-order cumulant of $s(t)$ defined by $\gamma_s := \text{cum}\{s(t), s(t), s^*(t), s^*(t)\}$, α_k denotes a positive value (in Section III-B, it will be shown how we choose the values of α_k's), the superscript $*$ denotes the complex conjugate, and σ_x^2 denotes the variance of $x(t)$, which is the output of the equalizer $W(z)$ (see Fig. 1).

The main difference between the two-step procedures in the conventional SEMs (e.g., [3], [5], [7], [9], [12]) and the one proposed here is the denominator of the first step, that is, the conventional first-step procedures include the second-order cumulants of $s(t)$, whereas our proposed one, that is, (6), possesses only higher order cumulants of $s(t)$.

As for the convergence of the two-step iterative procedure (6) and (7), under the assumptions that the equalizer $W(z)$ is an IIR filter and that the noise in (1) is absent, we show only a theorem of convergence (Theorem 1). The reader is referred to [9] for the proof. However, when we must take account of cases such that the equalizer is not of sufficient length for $H(z)$ (i.e., an undermodeled case shown in [8]) and that the noise has a strong power, we should note that the desired solutions of the two-step procedure may not fulfill (2) but may approximately fulfill (2). A complete analysis of the SEMs in undermodeled cases is shown in [8].

Theorem 1 [9]: Let $g^{(k)}(0)$ be an initial value for iterations of two steps (6) and (7) for each $k = -\infty, \cdots, \infty$. Let β_k be nonnegative scalar defined as

$$\beta_k = \left|\frac{1}{\alpha_k}\right|^{\frac{1}{p+q-1}}. \tag{8}$$

Let k_0 be $k_0 = \arg\max_{k\in\{-\infty,\cdots,\infty\}} \beta_k |g^{(k)}(0)|$. Suppose the index k_0 is unique, that is, $\beta_{k_0}|g^{(k_0)}(0)| > \beta_k |g^{(k)}(0)|$ for any other $k \in \{-\infty, \cdots, \infty\}$, then, as $i \to \infty$, it follows that

$$\lim_{i\to\infty}\left|g^{(k)}(i)\right| = \begin{cases} 0, & \text{for } k \neq k_0 \\ \tilde{c} \neq 0, & \text{for } k = k_0 \end{cases} \tag{9}$$

where $g^{(k)}(i)$ denotes the value obtained in the ith cycle of the iterations of two steps (6) and (7) and $\tilde{c}$ is a scalar positive constant.

Remark 1: It is shown in [5, Section IV] that the integer k_0 shown in Theorem 1 is unique except for pathological cases.

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

To find the solution $\hat{W}(z)$ in (2), we adjust the elements of the vector $\boldsymbol{w}$ so that $\boldsymbol{g} = \boldsymbol{H}\boldsymbol{w}$ is equal to the vector $\boldsymbol{\delta}^{(k_1)}$ whose nth element is $c\delta(n-k_1)$ for some fixed k_1, where $\delta(t)$ is the Kronecker delta function and k_1 is an integer standing for the same time shift as k_1 in (2). However, since $\boldsymbol{w}$ is of finite length, it may be only required that $\boldsymbol{w}$ is chosen to minimize the distance (norm) between $\boldsymbol{H}\boldsymbol{w}$ and $\boldsymbol{\delta}^{(k_1)}$. Hence, in order to derive an algorithm with respect to $\boldsymbol{w}$, we consider the following weighted least-squares problem:

$$\min_{\boldsymbol{w}} (\boldsymbol{H}\boldsymbol{w} - \boldsymbol{g})^{T*} \boldsymbol{\Lambda} (\boldsymbol{H}\boldsymbol{w} - \boldsymbol{g}). \tag{10}$$

Here, $\boldsymbol{\Lambda}$ is a diagonal matrix whose diagonal elements all are positive values. The solution is known to be given by

$$\boldsymbol{w} = (\boldsymbol{H}^{T*}\boldsymbol{\Lambda}\boldsymbol{H})^{-1}\boldsymbol{H}^{T*}\boldsymbol{\Lambda}\boldsymbol{g}. \tag{11}$$

Note that, from assumption A1), $\boldsymbol{H}^{T*}\boldsymbol{\Lambda}\boldsymbol{H}$ is invertible for any L, because $\boldsymbol{H}$ is of full column rank and $\boldsymbol{\Lambda}$ is a nonsingular diagonal matrix (this fact is also mentioned in [9, p. 508, line 10] without proof). The update rules of $\boldsymbol{w}$ in the conventional and the proposed SEMs are based on (11).

In the conventional SEMs [5], [7], [9], [12], the positive diagonal elements of $\boldsymbol{\Lambda}$ in (11) are set to 1 or the variance of the input $s(t)$. This means that $\boldsymbol{H}^{T*}\boldsymbol{\Lambda}\boldsymbol{H}$ is calculated by the second-order statistics of the output $y(t)$. We consider that this is the reason why the conventional SEMs are sensitive to Gaussian noise.

In what follows, we shall show that $\boldsymbol{H}^{T*}\boldsymbol{\Lambda}\boldsymbol{H}$ in (11) can be applied to a set of fourth-order cumulants of the output $y(t)$, if we choose appropriately a diagonal matrix $\boldsymbol{\Lambda}$ in (10). To this end, as the diagonal elements λ_k $(k = -\infty, \cdots, \infty)$ of $\boldsymbol{\Lambda}$, we choose the λ_k's expressed as

$$\lambda_k := \text{sign}(\gamma_s)\gamma_s\tilde{\alpha}_k, \qquad k = -\infty, \cdots, \infty \tag{12}$$

$$\tilde{\alpha}_k := \sum_{l=L_1}^{L_2} |h_{kl}|^2 \tag{13}$$

where $\text{sign}(\gamma)$ in (12) denotes the sign of γ, that is, $\text{sign}(\gamma) = 1$ if $\gamma > 0$, $\text{sign}(\gamma) = 0$ if $\gamma = 0$, and $\text{sign}(\gamma) = -1$ if $\gamma < 0$, and

h_{kl} in (13) denotes the element of $\boldsymbol{H}$ in (5), that is, the parameter $h^{(k-l)}$ of $H(z)$, $k = -\infty, \cdots, \infty, l = L_1, (L_1+1), \cdots, L_2$.

Remark 2: The matrix $\boldsymbol{\Lambda}$ is generally a nonsingular matrix except for pathological cases, and the elements of $\boldsymbol{\Lambda}$ are positive values. To avoid completely the pathological cases, the parameters L_1 and L_2 in (13) must be set to enough large negative and positive values (say, $-\infty$ and $+\infty$), respectively. Then $\tilde{\alpha}_k$ in (13) becomes a positive constant value for all k's.

From (12) and (13), $\boldsymbol{\Lambda}$ can be expressed as $\text{sign}(\gamma_s)\boldsymbol{I}\tilde{\boldsymbol{\Lambda}}$, where $\boldsymbol{I}$ is the identity matrix and $\tilde{\boldsymbol{\Lambda}}$ is also a diagonal matrix whose elements are $\gamma_s\tilde{\alpha}_k$, $k = -\infty, \cdots, \infty$. Then, substituting $\text{sign}(\gamma_s)\boldsymbol{I}\tilde{\boldsymbol{\Lambda}}$ into $\boldsymbol{\Lambda}$ in (11), the right-hand side of (11) becomes

$$(\boldsymbol{H}^{T*}\tilde{\boldsymbol{\Lambda}}\boldsymbol{H})^{-1}\boldsymbol{H}^{T*}\tilde{\boldsymbol{\Lambda}}\boldsymbol{g} \tag{14}$$

because $\text{sign}(\gamma_s)\boldsymbol{I}$ is a diagonal matrix whose elements all are either $+1$ or -1.

Here, $\boldsymbol{H}^{T*}\tilde{\boldsymbol{\Lambda}}\boldsymbol{H}$ in (14) can be expressed by the fourth-order cumulants matrix of $y(t)$, which is defined by $[\boldsymbol{C}_{y,l}^{(4)}]_{r_1,r_2} = \text{cum}\{y(t-r_1), y^*(t-r_2), y(t-l), y^*(t-l)\}$ [10], that is,

$$\boldsymbol{H}^{T*}\tilde{\boldsymbol{\Lambda}}\boldsymbol{H} := \sum_{l=L_1}^{L_2} \boldsymbol{C}_{y,l}^{(4)} \tag{15}$$

where $[X]_{r_1,r_2}$ denotes the (r_1, r_2)th element of the $L \times L$ matrix X, in which r_i's take the values of $L_1, (L_1+1), \cdots, L_2$. As for $\boldsymbol{H}^{T*}\tilde{\boldsymbol{\Lambda}}\boldsymbol{g}$ in (14), by using (6) with $\alpha_k = \tilde{\alpha}_k$ in (13) and the similar way as in [5], it can be given by

$$\boldsymbol{d} := [d_{L_1}, d_{L_1+1}, \cdots, d_{L_2}]^T \tag{16}$$

where d_l's are given by $d_l = \text{cum}\{\underbrace{x(t), \cdots, x(t)}_{p}, \underbrace{x^*(t), \cdots, x^*(t)}_{q}, y^*(t-j)\}$ $(l = L_1, L_1+1, \cdots, L_2)$. Therefore, it can be seen from (15) and (16) that the right-hand side of (11) can be calculated by the fourth-order statistics of the output $y(t)$, provided that $\boldsymbol{\Lambda}$ in (10) is replaced by $\text{sign}(\gamma_s)\boldsymbol{I}\tilde{\boldsymbol{\Lambda}}$. Then, (14) can be expressed as

$$\boldsymbol{w}^{[1]} = \boldsymbol{R}^{-1}\boldsymbol{d} \tag{17}$$

where $\boldsymbol{R} := \sum_{l=L_1}^{L_2} \boldsymbol{C}_{y,l}^{(4)}$. It can be easily shown that the second step (7) is expressed as

$$\boldsymbol{w}^{[2]} := \frac{\boldsymbol{w}^{[1]}}{\sqrt{\sigma_x^2}}. \tag{18}$$

Hence, (17) and (18) are our proposed two steps to modify $\boldsymbol{w}$.

From (17), it can be seen that, since the update procedure of $\boldsymbol{w}$ consists of only higher order cumulants of $y(t)$, then the two-step procedure (17) and (18) becomes less sensitive to Gaussian noise. [Note that, since (18) is only used to normalize $\boldsymbol{w}$, even if σ_x^2 is a second-order statistic, there is less effect of Gaussian noise for finding the desired solution $\hat{\boldsymbol{w}}$, that is, $\boldsymbol{H}\hat{\boldsymbol{w}} = \boldsymbol{\delta}^{(k_1)}$.] This is a novel key point of our proposed SEM, from which the proposed method is referred to as a *robust super-exponential method* (RSEM).

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 $H(z)$ was set to be an FIR filter of length 7 with the impulse responses (0.4, 1, -0.7, 0.6, 0.3, -0.4, 0.1), which is the same system as in [9]. We used an equalizer of length $L = 16$ which was initialized to $\boldsymbol{w}(0) = [0,0,0,0,0,1,0,0,0,0,0,0,0,0,0,0]^T$, which is also the same situation as in [9]. The input $s(t)$ of the system $H(z)$ was sub-Gaussian which takes one of two values, -1 and 1 with equal probability 1/2. The parameters p and q in (6) were set to be $p = 2$ and $q = 1$, respectively, that is, κ_s in (3) was the fourth-order cumulants of $s(t)$. Then the value of κ_s is -2. The Gaussian noise $n(t)$ with its variance σ_n^2 was included in the output $y(t)$ at various SNR levels. The SNR was considered at the output of the system $H(z)$.

The matrix $\boldsymbol{R}$ in (17) was calculated using a moving average defined by

$$\begin{aligned}\boldsymbol{R}(t) := \beta_1\boldsymbol{R}(t-1) + (1-\beta_1)\Big\{&\boldsymbol{V}_1(t)\boldsymbol{V}_1^{T*}(t) \\ &-\boldsymbol{V}_1(t)\tilde{\boldsymbol{V}}_1^{T*}(t) - \boldsymbol{V}_2(t)\tilde{\boldsymbol{V}}_2^{T*}(t) - \text{tr}\Big\{\tilde{\boldsymbol{V}}_1(t)\Big\}\boldsymbol{V}_1(t)\Big\}\end{aligned} \tag{19}$$

in which $\boldsymbol{V}_1(t) := \boldsymbol{y}(t)\boldsymbol{y}^{T*}(t)$, $\boldsymbol{V}_2(t) := \boldsymbol{y}(t)\boldsymbol{y}^T(t)$, $\text{tr}\{X\}$ denotes the trace of the matrix X, and $\tilde{\boldsymbol{V}}_i(t)$ is a moving average of $\boldsymbol{V}_i(t)$ calculated by

$$\tilde{\boldsymbol{V}}_i(t) = \beta_2\tilde{\boldsymbol{V}}_i(t-1) + (1-\beta_2)\boldsymbol{V}_i(t), \quad i = 1, 2. \tag{20}$$

Here, $\boldsymbol{y}(t)$ denotes an L-column vector whose elements are $y(t-L_1), y(t-L_1-1), \cdots, y(t-L_2)$. The vector $\boldsymbol{d}$ in (17) was calculated using a moving average defined by

$$\begin{aligned}\boldsymbol{d}(t) := \beta_1\boldsymbol{d}(t-1) + (1-\beta_1)\{&|x(t)|^2x(t)\boldsymbol{y}^*(t) \\ &-2\tilde{v}_{x1}(t)x(t)\boldsymbol{y}^*(t) - \tilde{v}_{x2}(t)x^*(t)\boldsymbol{y}^*(t)\}\end{aligned} \tag{21}$$

in which $\tilde{v}_{xi}(t)$ is a moving average defined by

$$\tilde{v}_{xi}(t) = \beta_2\tilde{v}_{xi}(t-1) + (1-\beta_2)v_{xi}(t), \quad i = 1, 2 \tag{22}$$

where $v_{x1}(t) = x(t)x^*(t)$ and $v_{x2}(t) = x(t)^2$. The parameters β_1 and β_2 were set to be 0.999 99 and 0.995, respectively.

As a measure of performance, we used the intersymbol interference (ISI) defined in the logarithmic (dB) scale by

$$\text{ISI} = 10\log_{10}\left\{\frac{\left(\sum_k \left|g^{(k)}\right|^2 - |g_{\max}|^2\right)}{|g_{\max}|^2}\right\} \tag{23}$$

where $g_{\max}$ is the component of $g^{(k)}$ having the maximal absolute value (the leading tap). The value of ISI becomes $-\infty$ if $g^{(k)}$'s satisfying (2) are obtained and, hence, a large negative value of ISI indicates the proximity to the desired solution. However, the ISI is not enough as a measure of performances. Thus, we used the mean-squared estimation error (MSE) defined as

$$\text{MSE} = E\left[(\tilde{x}(t) - s(t-k_1))^2\right] \tag{24}$$

and the bit error rate (BER), where $E[x]$ in (24) denotes the expectation of x, $\tilde{x}(t) = \xi x(t)$, in which the parameter ξ is $1/g_{\max}$ so that the leading tap $g_{\max}$ becomes 1, and k_1 denotes a constant delay. We note that $g_{\max}$ and k_1 are found when the ISI in (23) is calculated. When the BER is calculated, we apply $\text{sign}(x(t))$ to the output of the equalizer $x(t)$.

For comparison, the SEM proposed in [9] was used. The parameters β_1 in (21) and β_2 in (22) were set to be 0.9998 and 0.999, respectively. In the SEM, $\boldsymbol{R}$ in (17) was calculated by $\tilde{\boldsymbol{V}}_1(t)$ in (20), where the parameter β_2 in (20) was set to be 0.999.

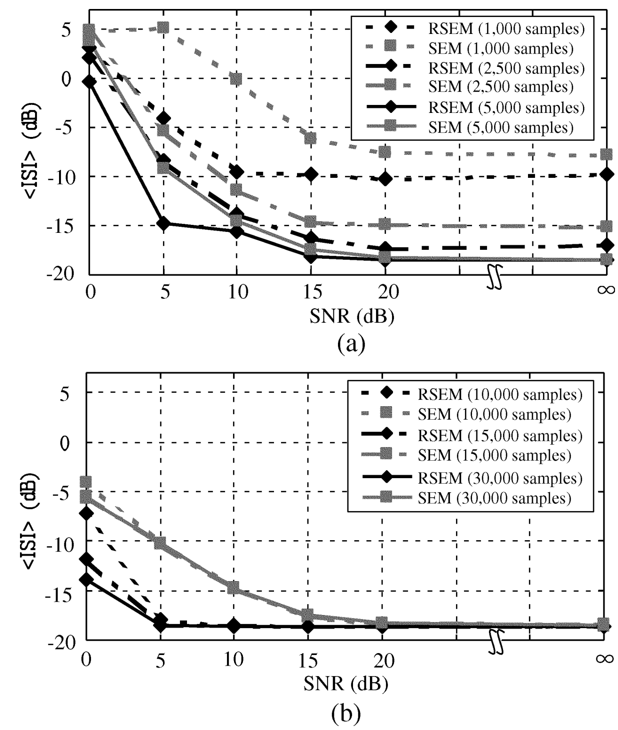

Fig. 2. Performances of the RSEM and the SEM with varying SNR levels (a) in the cases of 1000, 2500, and 5000 samples and (b) in the cases of 10 000, 15 000, and 30 000 samples.

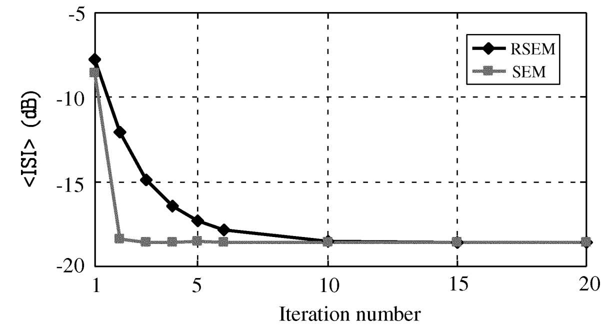

Fig. 3. Performances of the RSEM and the SEM with varying number of iterations in the cases that the SNR level is ∞ and the data length is 10 000 samples.

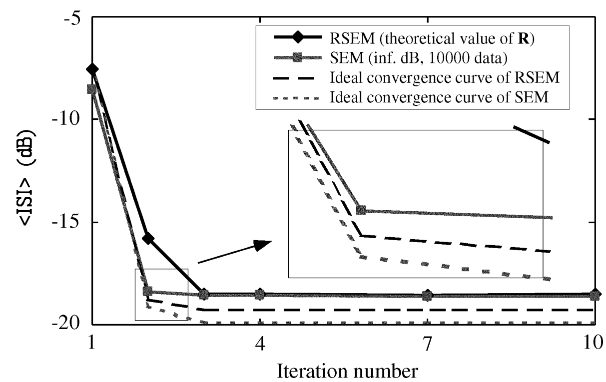

Fig. 4. Performances of the RSEM for the cases that $\boldsymbol{R}$ is its theoretical value and the SEM for $\mathrm{SNR} = \infty$ dB (10 000 samples) with varying number of iterations, and the ideal convergence curves of the RSEM and the SEM.

Fig. 2 shows the results of the ISIs of the RSEM and SEM, in the cases where the SNR levels were taken to the values ranging from 0 through 20 dB to ∞ dB ($\sigma_n^2 = 0$). The ISIs shown in Fig. 2 are the average of the results obtained by 100 Monte Carlo trials, and, for each trial, the vector $\boldsymbol{w}$ was modified by 20 iterations; for each iteration, the matrix $\boldsymbol{R}$ corresponding to each method and the vector $\boldsymbol{d}$ were estimated with (19)–(22), using several data samples (see Fig. 2). The vertical and horizontal axes in Fig. 2 represent the average of ISIs denoted by $\langle \mathrm{ISI} \rangle$ and the SNR, respectively.

It can be seen from Fig. 2(a) that, when the data samples are less than 5000, for each sample data and each SNR level, the performances of the RSEM after 20 iterations are better than those of the SEM. However, in order to use the feature of the RSEM, which is insensitive to Gaussian noise, at least 5000 samples are needed. Note that this statement can be effective for the case that the SNR level is more than 5 dB. From Fig. 2(b), it can be seen that, as the number of data samples increases, the RSEM gives better performance for every SNR level, whereas the performance of the SEM hardly changes. However, if the obtained results are viewed from the point of convergence speed, it can be discovered that, in the SNR levels of more than 15 dB, the convergence speed of the RSEM is slower than the SEM (see Fig. 3). [Note that Fig. 3 shows the case of 10 000 samples and $\mathrm{SNR} = \infty$ dB. The horizontal axis in Fig. 3 denotes the number of iterations.] This tendency was observed when the number of data samples is 5000 and the SNR level is more than 10 dB. This has resulted from the fact that it is difficult to estimate the fourth-order cumulant matrix $\boldsymbol{R}$ in the RSEM with high accuracy, compared with the second-order cumulant matrix $\boldsymbol{R}$ in the SEM. However, even if the matrix $\boldsymbol{R}$ corresponding to the RSEM is its theoretical value, the RSEM has a slightly slower convergence than the SEM (see Fig. 4). Indeed, it can be seen from Fig. 4 that the convergence ability of the RSEM is slightly worse than that of the SEM (see the dashed and dotted lines, respectively, in Fig. 4, which were obtained by using the theoretical values of $\boldsymbol{R}$ and $\boldsymbol{d}$ corresponding to each method).

From these results, we consider that the RSEM is effective for solving the blind equalization problem in the presence of Gaussian noise. However, if quick responses are demanded to the equalizers in a case where, for 5000 samples the SNR level is more than 10 dB or for more than 10 000 sample, the SNR level is more than 15 dB, it is better to use the SEM. For less than 2500 samples, we consider that one can choose either the RSEM or the SEM.

Fig. 5 shows the results of the MSEs and bit error rates (BERs) which are the averages of the results obtained by 100 Monte Carlo trials. For each trial, the vector $\boldsymbol{w}$ was modified for 20 iterations, using (17) and (18) corresponding to each method, and the MSE and BER was calculated by using 1000 outputs of the equalizer with the modified $\boldsymbol{w}$. For each iteration, the matrix $\boldsymbol{R}$ corresponding to each method and the vector $\boldsymbol{d}$ were estimated by 15 000 samples, using (19)–(22). Since the values of the BERs of both the methods in the cases of $\mathrm{SNR} = 20$ and ∞ dB were zero, these values are not plotted in Fig. 5.

From Fig. 5, it can be seen that, as the SNR level is lower, the MSE of the RSEM becomes worse than that of the SEM. This results from the fact that the convergence point of the SEM is essentially equivalent to Wiener filtering such that the MSE is minimized. Under the influence of the results, the BER of the RSEM is also worse than that of the SEM. However, it can be seen that, although the MSE of the RSEM becomes worse as the SNR level is lower, the difference between the BERs of the two methods becomes smaller. This results from the fact that the

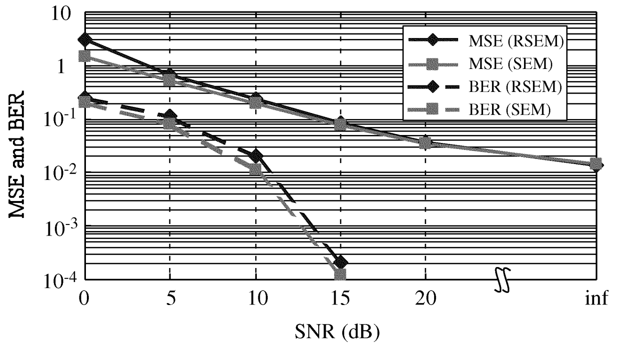

Fig. 5. MSEs and BERs of the RSEM and the SEM with varying SNR levels.

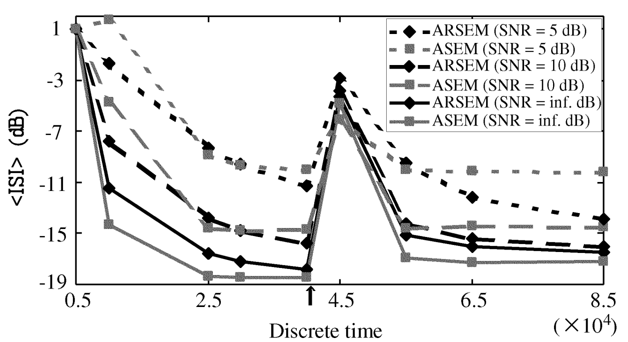

Fig. 6. Performances obtained by adaptively implementing the RSEM and the SEM, in each of the three cases, $\mathrm{SNR} = 5$ dB, 10 dB, and ∞ dB.

TABLE I
FLOPS AND ELAPSED TIME OF THE RSEM AND SME FOR ONE ITERATION

Method	Flops ($\log_{10}$ FLO)	Time (sec.)
RSEM	7.143	0.0893
SEM	5.856	0.0395

RSEM can achieve estimating the inverse of $H(z)$ with as little influence of Gaussian noise as possible. From this result and this property of the RSEM, we motivate to the investigation of estimating the inputs by using a Kalman filter (KF), for example, the KF proposed by Marcos [6]. However, this is beyond the scope of this paper. Therefore, the results will be shown in a forthcoming paper.

Fig. 6 shows the results of the performances obtained by adaptively implementing the RSEM and SEM, in the cases of $\mathrm{SNR} = 5$, 10, and 20 dB. In this example, the vector $\boldsymbol{w}$ is adaptively modified by using

$$\boldsymbol{w}^{[1]}(t) = \boldsymbol{R}^{-1}(t-1)\boldsymbol{d}(t-1) \tag{25}$$

$$\boldsymbol{w}(t) = \frac{\boldsymbol{w}^{[1]}(t)}{\sqrt{\sigma_x^2}} \tag{26}$$

where the vector $\boldsymbol{w}^{[1]}(t)$ on the right-hand side of (26) is equal to $\boldsymbol{w}^{[1]}(t)$ on the left-hand of (25). At $t = 40\,001$ (see $\uparrow$ in Fig. 6), the impulse response $h^{(3)} = 0.3$ of $H(z)$ was changed to $h^{(3)} = -0.3$. Note that for $t = 1$ through 5,000 and 40 001 through 45,000, $\boldsymbol{w}$ was not modified, but $\boldsymbol{R}$ and $\boldsymbol{d}$ were estimated by using (19)–(22).

It can be seen from Fig. 6 that, as the SNR decreases, the adaptive RSEM (ARSEM) appears to converge to better performance, compared with the adaptive SEM (ASEM). A detailed investigation of the performances of the ARSEM and ASEM will be dealt with in a forthcoming paper, because of the page limitation.

Finally, we shall show the computational complexities of the RSEM and SEM by using the flops (floating-point operations) and elapsed time for one iteration for 5000 samples. Table I shows these results. These values pertain to a 3.06-GHz–1.00-GB machine. For the interested reader, these values of other blind equalization algorithm (including the SEM) can be found in [2].

From all of the results, we conclude that, although the computational complexity of the RSEM is more than the SEM, the RSEM indicates better performances than the SEM in the case that the SNR level is lower than 10 dB; hence, for solving the blind equalization in the presence of Gaussian noise, the RSEM has sufficient merits.

V. Conclusion

We have proposed an SEM for solving a blind equalization problem, which is referred to as a *robust super-exponential method* (RSEM). The RSEM is robust against Gaussian noise, which means that the RSEM can be used to estimate the inverse of the unknown transfer function $H(z)$, even if Gaussian noise is added to the output of $H(z)$ [see (1)]. This is a novel property of the proposed method not possessed by conventional SEMs. Using the computer simulations, we presented several results in which the RSEM was compared with the SEM under several conditions.

As for further work, we consider extending the RSEM so that the RSEM can be applied to multiple-input multiple-output systems. In this research, we will deal with the RSEMs incorporated with Kalman filters and their adaptive implementations besides.

References

[1] Z. Ding and G. Li, “Single-channel blind equalization for GSM cellular systems,” *IEEE J. Sel. Areas Commun.*, vol. 16, no. 8, pp. 1493–1505, Oct. 1998.

[2] S. Fiori, “Fast fixed-point neural blind-deconvolution algorithm,” *IEEE Trans. Neural Netw.*, vol. 15, no. 2, pp. 455–459, Mar. 2004.

[3] F. Herrmann and A. K. Nandi, “Reduced computation blind super-exponential equaliser,” *Electron. Lett.*, vol. 34, no. 23, pp. 2208–2209, Nov. 1998.

[4] Y. Inouye, “Autoregressive model fitting for multichannel time series of degenerate rank: limit properties,” *IEEE Trans. Circuits Syst.*, vol. CAS-32, no. 3, pp. 252–259, Mar. 1985.

[5] 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.

[6] S. Marcos, “A network of adaptive Kalman filters for data channel equalization,” *IEEE Trans. Signal Process.*, vol. 48, no. 9, pp. 2620–2627, Sep. 2000.

[7] M. Martone, “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.

[8] P. A. Regalia and M. Mboup, “Undermodeled equalization: A characterization of stationary points for a family of blind criteria,” *IEEE Trans. Signal Process.*, vol. 47, no. 3, pp. 760–770, Mar. 1999.

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

[10] 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.

[11] K. Yang, T. Ohira, Y. Zhang, and C.-Y. Chi, “Super-exponential blind adaptive beamforming,” *IEEE Trans. Signal Process.*, vol. 52, no. 6, pp. 1549–1563, Jun. 2004.

[12] 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.