ROBUST SUPER-EXPONENTIAL METHODS FOR BLIND EQUALIZATION OF AN SISO-IIR SYSTEM WITH ADDITIVE GAUSSIAN NOISE

Mitsuru Kawamoto^a, Kiyotaka Kohno^b, Yujiro Inouye^b, and Asoke K. Nandi^c

a. Information Technology Research Institute,
National Institute of Advanced Industrial Science and Technology,
Central 2, 1-1-1 Umezono, Tsukuba, Ibaraki 305-8568, JAPAN
b. Department of Electronic and Control Systems Engineering, Shimane University,
1060 Nishikawatsu, Matsue, Shimane 690-8504, Japan
c. Department of Electrical Engineering and Electronics, The University of Liverpool,
Brownlow Hill, Liverpool, L69 3GJ, UK

ABSTRACT

The present 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.

1. 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 SISO systems, which is called the super-exponential method (SEM) [6], and then several researchers extended the idea of the SEM, e.g., see [2, 4, 8]. and references therein. One of the attractive properties of the SEMs 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 [6].

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
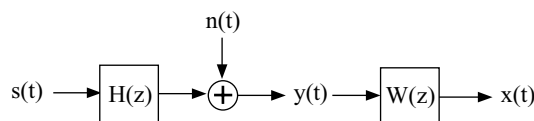

Fig. 1. The composite system of an unknown system and a filter.

robust super-exponential method (RSEM). Computer simulations are presented to demonstrate the validity of the proposed RSEM.

2. PROBLEM FORMULATION AND ASSUMPTIONS

We consider a single-input single-output (SISO) system with an additive noise as described by

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

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 $H(z) = \sum_{k=-\infty}^{\infty} h^{(k)} z^k$, $y(t)$ 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 $W(z) = \sum_{k=-\infty}^{\infty} w^{(k)} z^k$ so that $G(z) := W(z)H(z)$ becomes

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

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. 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.

427

A1) The unknown system $H(z)$ is a stable, possibly non-minimum 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

$$\kappa_s = \text{cum}\{\underbrace{s(t), \cdots, s(t)}_{p}, \underbrace{s^*(t), \cdots, s^*(t)}_{q+1}\}, \qquad (3)$$

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 an FIR system of sufficient length $L = L_2 - L_1 + 1$ so that the truncation effect can be ignored.

The assumption A3) is necessary for an easy implementation of the equalizer $W(z)$. However, in order to ensure the convergence of the proposed method presented in the next section, the equalizer should be an IIR system (see Theorem 1). Therefore, the assumption A3) ensures an easy implementation of the equalizer, but only an approximate convergence of the proposed method.

The combined system response subject to the finite length restriction is

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

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

$$\boldsymbol{g} = \boldsymbol{H}\boldsymbol{w}, \qquad (5)$$

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 $h_{kl} = h^{(k-l)}$, $k = -\infty, \cdots, \infty$, $l = L_1, (L_1+1), \cdots, L_2$.

3. ROBUST SUPER-EXPONENTIAL METHODS (RSEMS)

3.1. Two-step iterative procedure of vector 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:

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

$$g^{(k)[2]} = g^{(k)[1]}/\sqrt{\sigma_x^2}, \qquad (7)$$

where $(\cdot)^{[1]}$, $(\cdot)^{[2]}$ 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 subsection 3.2, 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., [2, 4, 6, 8]) and the proposed one 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 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 [6] 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 [5], in this paper, we consider an undermodeled case) 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 [5].

Theorem 1 [6]: 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 non-negative scalar defined as

$$\beta_k = |1/\alpha_k|^{\frac{1}{p+q-1}} \qquad (8)$$

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

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

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 [2, Section IV] that the integer k_0 shown in Theorem 1 is unique except for pathological cases.*

3.2. Two-step iterative procedure for 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:

$$\min_{\boldsymbol{w}} (\boldsymbol{H}\boldsymbol{w} - \boldsymbol{g})^{T*} \boldsymbol{\Lambda} (\boldsymbol{H}\boldsymbol{w} - \boldsymbol{g}). \qquad (9)$$

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

$$\boldsymbol{w} = (\boldsymbol{H}^{T*} \boldsymbol{\Lambda} \boldsymbol{H})^{-1} \boldsymbol{H}^{T*} \boldsymbol{\Lambda} \boldsymbol{g}. \qquad (10)$$

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 non-singular diagonal matrix (this fact is also mentioned in [6, p. 508, line 10] without proof). The update rules of $\boldsymbol{w}$ in the conventional and the proposed SEMs are based on (10).

In the conventional SEMs ([2, 4, 6, 8]), the positive diagonal elements of $\boldsymbol{\Lambda}$ in (10) 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 (10) 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 (9). To this end, as the diagonal elements λ_k ($k = -\infty, \cdots, \infty$) of $\boldsymbol{\Lambda}$, we choose the λ_k's expressed as

$$\lambda_k := \text{sign}(\gamma_s)\gamma_s\tilde{\alpha}_k, \qquad k = -\infty, \cdots, \infty, \quad (11)$$

$$\tilde{\alpha}_k := \sum_{l=L_1}^{L_2} |h_{kl}|^2, \qquad (12)$$

where $\text{sign}(\gamma)$ in (11) 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 (12) 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 (12) must be set to enough large negative and positive values (say, $-\infty$ and $+\infty$), respectively. Then $\tilde{\alpha}_k$ in (12) becomes a positive constant value for every k's.*

From (11) and (12), $\boldsymbol{\Lambda}$ 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 (10), the right-hand side of (10) becomes

$$(\boldsymbol{H}^{T*}\tilde{\boldsymbol{\Lambda}}\boldsymbol{H})^{-1}\boldsymbol{H}^{T*}\tilde{\boldsymbol{\Lambda}}\boldsymbol{g}, \qquad (13)$$

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

Here, $\boldsymbol{H}^{T*}\tilde{\boldsymbol{\Lambda}}\boldsymbol{H}$ in (13) can be expressed by the fourth-order cumulants matrix of $y(t)$, which is defined by $[\boldsymbol{C}_{y,l}^{(4)}]_{r_1,r_2} = \text{cum}\{y(t-r_1), y^*(t-r_2), y(t-l), y^*(t-l)\}$ [7], that is,

$$\boldsymbol{H}^{T*}\tilde{\boldsymbol{\Lambda}}\boldsymbol{H} := \sum_{l=L_1}^{L_2} \boldsymbol{C}_{y,l}^{(4)}, \qquad (14)$$

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 (13), by using (6) with $\alpha_k = \tilde{\alpha}_k$ in (12) and the similar way as in [2], it can be given by

$$\boldsymbol{d} := [d_{L_1}, d_{L_1+1}, \cdots, d_{L_2}]^T, \qquad (15)$$

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 (14) and (15) that the right-hand side of (10) can be calculated by the fourth-order statistics of the output $y(t)$, provided that $\boldsymbol{\Lambda}$ in (9) is replaced by $\text{sign}(\gamma_s)\boldsymbol{I}\tilde{\boldsymbol{\Lambda}}$. Then, (13) can be expressed as

$$\boldsymbol{w}^{[1]} = \boldsymbol{R}^{-1}\boldsymbol{d}, \qquad (16)$$

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

$$\boldsymbol{w}^{[2]} := \boldsymbol{w}^{[1]}/\sqrt{\sigma_x^2}. \qquad (17)$$

Hence, (16) and (17) are our proposed two steps to modify $\boldsymbol{w}$.

From (16), 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 (16) and (17) becomes less sensitive to Gaussian noise. [Note that since (17) 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).

4. COMPUTER SIMULATIONS

To demonstrate the validity of the proposed method, many computer simulations were conducted. Some results are shown in this section. The simulation conditions, for example, the settings of the unknown system $H(z)$ and the equalizer $W(z)$, the input signal $s(t)$, the parameters p and q in (6), the definition of SNR, etc, were the same as those in [3]. On the details of the simulation conditions, see [3].

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

$$\text{ISI} = 10\log_{10}\{(\sum_k |g^{(k)}|^2 - |g_{\max}|^2)/|g_{\max}|^2\} \quad (18)$$

where $g_{\max}$ is the component of $g^{(k)}$ having the maximal absolute value (the leading tap). However, the ISI is not enough as a measure of performances. Thus by using the mean-squared estimation error (MSE) and the bit error rate (BER), we measured the performances of the RSEM. However, the results are not put in the paper, because of the page limitation. For the interested reader, the results will be shown in [3]. For comparison, the SEM proposed in [6] was used. On the simulation conditions of the SEM, see [3].

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 to 20 dB, and ∞ 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, and for each iteration, the matrix $\boldsymbol{R}$ corresponding to each method and the vector $\boldsymbol{d}$ were estimated, using several data samples (see Fig. 2). The vertical and horizontal axes in Fig. 2 represent the average of ISIs denoted by <ISI> and the SNR, respectively.
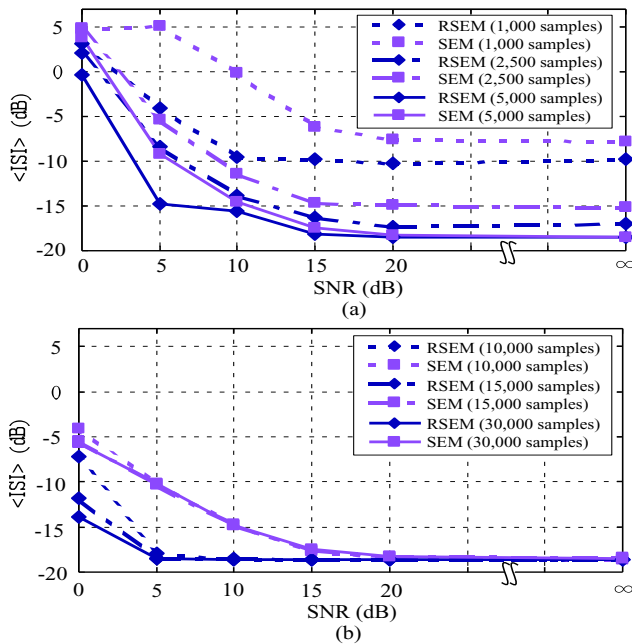
Fig. 2. The performances of the RSEM and the SEM with varying SNR levels, (a) in the cases of 1000 samples, 2500 samples, 5000 samples, (b) in the cases of 10000 samples, 15000 samples, 30000 samples.

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 5,000 samples are needed. Note that this data length is available 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 levels, whereas the performance of the SEM hardly changes.

However, if these results are viewed from the point of convergence speed, it can be discovered that there exists the case that the convergence speed of the RSEM is slower than the SEM (see Fig. 3). On the details of the results, see [3].

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, there exists the case that
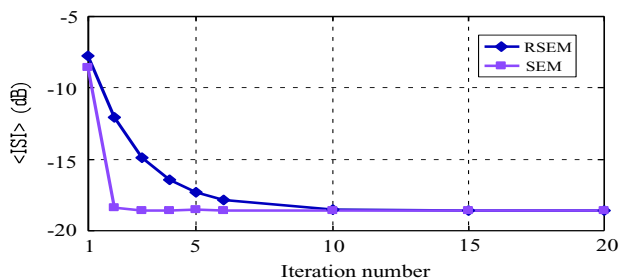
it is better to use the SEM.

5. CONCLUSIONS

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 the 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 (MIMO) systems.

6. REFERENCES

[1] Z. Ding and G. Li, "Single-channel blind equalization for GSM cellular systems," IEEE J. on Select. Areas in Commun., vol. 16, no. 8, Oct. 1998.

[2] 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.

[3] M. Kawamoto, M. Ohata, K. Kohno, Y. Inouye, and A.K. Nandi, "Robust super-exponential methods for blind equalization in the presence of Gaussian noise," to appear in IEEE Trans. Circuits and Systems II. (http://staff.aist.go.jp/m.kawamoto/publish.htm)

[4] M. Martone, "Fast adaptive super-exponential multistage beamforming cellular base-station transceivers with antenna arrays," IEEE Trans. Vehicular Tech., vol. 48, no. 4, pp. 1017-1028, Jul. 1999.

[5] P. A. Regalia and M. Mboup, "Undermodeled equalization: A characterization of stationary points for a family of blind criteria," IEEE Trans. Signal Processing, vol. 47, no. 3, pp. 760-770, Mar. 1999.

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

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

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

Fig. 3. The 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 10000 samples.