Learning Characteristics for General Class of Adaptive Blind Equalizer

Ye Li and K. J. Ray Liu
Electrical Engineering Department and Institute for Systems Research
University of Maryland, College Park, MD 20742
E-mail: liye(kjrliu)@src.umd.edu, Fax: (301) 314-9281

Abstract— **This paper presents a theoretical analysis of the static and dynamic convergence behavior for a general class of adaptive blind equalizers. We first study the properties of prediction error functions of blind equalization algorithms, and then we use these properties to analyze the static and dynamic convergence behavior based on the independence assumption. We prove in this paper that with a small step-size, the ensemble average of equalizer coefficients will converge to the minimum of the cost function near the channel inverse. However, the convergence is not consistent. The correlation matrix of equalizer coefficients at equilibrium is determined by a Lyapunov equation. According to our analysis results, for a given channel and step-size, there is an optimal length for an equalizer to minimize the intersymbol interference. This result implies that a longer-length blind equalizer does not necessarily outperform a shorter one, as contrary to what is conventionally conjectured. The theoretical analysis results are confirmed by computer simulations.**

I. Introduction

Since the pineering work by Sato[14], many blind channel equalization algorithms have been proposed[1], [2], [6], [17], [18]. They have been effectively used in digital communication systems to cancel the inter-symbol interference (ISI). Blind equalization algorithms are usually designed to minimize some cost functions consisting of higher-order statistics of the channel output, without using the channel input. They are implemented mostly by stochastic gradient algorithms. The convergence analysis of blind equalization algorithms is very important to understanding their performance.

Unlike most of the previous convergence analysis works[3], [7], [8], [10], [12], [15], [20], [21] which specifically focused on some blind equalization algorithm, we will present the static and dynamic convergence analysis for almost all adaptive blind equalization algorithms. Since there are many initialization strategies[4], [5], [9] to make blind equalizer reach an open eye pattern, we will concentrate on the convergence analysis when the coefficient sets of equalizers are near the global minima of their cost functions. In the static analysis, we derive the close form solution for the coefficients of FIR blind equalizers, from which we can evaluate the distortion caused by the finite length effect. In the dynamic analysis, we only use the *independence assumption*[7], which is widely used in the dynamic convergence analysis of adaptive algorithms [7], [10], [11], [12], [15], [20], [21]. Based on the independence assumption, together with the first-order approximation, we study the convergence and consistence of the equalizer coefficients. Our analysis indicates that for a given channel and step-size, there is an optimum length of equalizer minimizing the intersymbol interference, which implies that a longer blind equalizer does not necessarily perform better than a shorter one. This result can be applied to the design of blind equalizers used in digital communication systems.

The work was supported in part by the NSF grants MIP9309506 and MIP9457397.

II. Adaptive Blind Equalizers

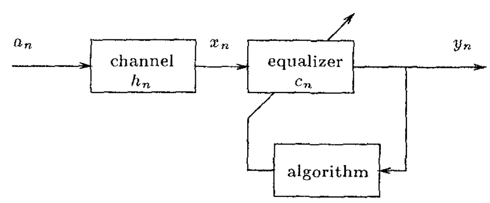

Fig. 1. PAM communication system with blind channel equalizer.

Without lose of generality, we consider a baseband representation of the pulse-amplitude-modulation (PAM) communication system with blind channel equalizer as shown in Figure 1. A sequence of independent, identically distributed (i.i.d.) digital signal $\{a_n \in \mathcal{R}\}$ with zero-mean and variance σ^2 is sent through a bounded-input bounded-output (BIBO) channel exhibiting linear distortion. The resulting output signal x_n can be expressed as

$$x_n = \sum_{k=-\infty}^{+\infty} a_k h_{n-k} + w_n, \tag{1}$$

where h_n is the impulse response of the linear time-invariant (LTI) channel, and w_n is white Gaussian channel noise. In this paper, we will ignore the effects of the channel noise.

As shown in Figure 1, a linear channel equalizer with parameters $\{c_n\}$ is used to remove the intersymbol interference caused by the channel distortion. The parameters $\{c_n\}$ are subject to adaptation via some algorithm to be determined. The equalizer output in Figure 1 can then be written as

$$\begin{aligned} y_n &= \sum_{k=-\infty}^{\infty} c_k x_{n-k} \\ &= \sum_{k=-\infty}^{\infty} s_k a_{n-k}, \end{aligned} \tag{2}$$

0-7803-3250-4/96$5.00©1996 IEEE

where $\{s_n\}$ is the impulse response of the equalized system related to h_n and c_n by

$$s_n \triangleq \sum_k h_k c_{n-k}. \tag{3}$$

In blind equalization, the original sequence is unknown to the receiver except for its probabilistic or statistical properties. A blind equalization algorithm is usually devised by minimizing a cost function consisting of the statistics of the output of the equalizer y_n, which is a function of $\{\cdots, s_{-1}, s_0, s_1, \cdots\}$ or $\{\cdots, c_{-1}, c_0, c_1, \cdots\}$, The cost function is usually of the form $E\{\Phi(y_n)\}$, where $\Phi(y_n)$, a function of y_n, is selected such that the cost function has the global minimum points at

$$\{s_n\} = \pm\{\delta[n-n_d]\} \text{ for all } n_d = 0, \pm 1, \pm 2, \cdots \tag{4}$$

A stochastic gradient algorithm is used to minimize the cost function to obtain an on-line equalization algorithm, which adjusts the k-th parameter of the equalizer at time n by

$$\widehat{c}_k^{(n+1)} = \widehat{c}_k^{(n)} - \mu\phi(y_n)x_{n-k}, \tag{5}$$

where μ is a small step size, $\phi(.)$ is the derivative of $\Phi(y_n)$, that is,

$$\phi(y_n) = \Phi'(y_n), \tag{6}$$

and it is sometimes called *prediction error function.*

If an FIR filter is used as the equalizer, then (5) can be expressed as

$$\widehat{\mathbf{c}}^{(n+1)} = \widehat{\mathbf{c}}^{(n)} - \mu\mathbf{x}_n\phi(y_n), \tag{7}$$

where $\widehat{\mathbf{c}}^{(n)}$ is the *coefficient vector* of a blind equalizer after n-th iteration defined as

$$\widehat{\mathbf{c}}^{(n)} \triangleq [\widehat{c}_{-N}^{(n)}, \cdots, \widehat{c}_0^{(n)}, \cdots, \widehat{c}_N^{(n)}]^T \tag{8}$$

and $\mathbf{x}_n$ is the *channel output vector* at time n defined as

$$\mathbf{x}_n \triangleq [x_{n+N}, \cdots, x_n, \cdots, x_{n-N}]^T. \tag{9}$$

Since all BIBO channels can be approximated as a moving-average model with appropriate impulse response $\{h_{-M}, \cdots, h_0, \cdots, h_M\}$, the channel output vector can be expressed as

$$\mathbf{x}_n = \mathcal{H}^T\mathbf{a}_n, \tag{10}$$

where $\mathcal{H}$ is a $(2N+2M+1) \times (2N+1)$ *channel matrix* defined as

$$\mathcal{H} \triangleq \begin{pmatrix} h_{-M} & 0 & \cdots & 0 \\ \vdots & h_{-M} & \ddots & \vdots \\ h_0 & \vdots & \ddots & 0 \\ \vdots & h_0 & \ddots & h_{-M} \\ h_M & \vdots & \ddots & \vdots \\ 0 & h_M & \ddots & h_0 \\ \vdots & 0 & \ddots & \vdots \\ 0 & \cdots & \cdots & h_M \end{pmatrix}, \tag{11}$$

and $\mathbf{a}_n$ is the *input symbol vector* at time n defined as

$$\mathbf{a}_n \triangleq [a_{n+(N+M)}, \cdots, a_n, \cdots, a_{n-(N+M)}]^T. \tag{12}$$

With the above definitions, the channel output can be expressed in a compact form:

$$\begin{aligned} y_n &= \mathbf{a}_n^T\widehat{\mathbf{s}}^{(n)} \\ &= \mathbf{a}_n^T\mathcal{H}\widehat{\mathbf{c}}^{(n)}, \end{aligned} \tag{13}$$

where $\widehat{\mathbf{s}}^{(n)}$ is the equalized system vector at time n defined as

$$\widehat{\mathbf{s}}^{(n)} \triangleq \mathcal{H}\widehat{\mathbf{c}}^{(n)}. \tag{14}$$

It is obvious that an FIR channel can not be perfectly equalized by an FIR equalizer, that is, there is no equalization vector $\mathbf{c}$ such that

$$\mathcal{H}\mathbf{c} = \mathbf{e}_{M+N}, \tag{15}$$

where

$$\mathbf{e}_{M+N} = [\underbrace{0, \cdots, 0}_{M+N},\ 1,\ \underbrace{0, \cdots, 0}_{M+N}]^T. \tag{16}$$

However, when the length of the equalizer is large enough, there exists a $\tilde{\mathbf{c}}$ such that $\| \mathcal{H}\tilde{\mathbf{c}} - \mathbf{e}_{M+N} \|$ is very small.

III. Properties of Prediction Error Function

Before analyzing the convergence behavior of blind equalizers, we first introduce some properties of the prediction error function here. The following lemma considers two important properties to be used in subsequent discussions.

Lemma: *The prediction error function $\phi(.)$ has the following two properties:*

1) *When the parameters of a finite-length equalizer make its cost function attain one of its minima, the output of the equalized system, $\tilde{y}_n$, satisfies*
(i) $E\{\phi(\tilde{y}_n)\mathbf{x}_n\} = \mathbf{0}$, and
(ii) $\mathcal{H}^T\tilde{F}\mathcal{H}$ is positive-definite,
where the $(2M+2N+1) \times (2M+2N+1)$ matrix $\tilde{F}$ is defined as

$$\tilde{F} = \frac{1}{\sigma^2}E\{\mathbf{a}_n\phi'(\tilde{y}_n)\mathbf{a}_n^T\}, \tag{17}$$

with $\phi'(.)$ being the derivative of $\phi(.)$, $\tilde{y}_n = \sum_k \tilde{c}_k x_{n-k}$ and $\tilde{c}_k$ being the equalizer coefficients making the cost function attain a minimum.

2) *For all integers n and k*

$$E\{\phi(a_n)a_k\} = 0, \tag{18}$$

and

$$E\{\phi'(a_n)a_k^2\} > 0. \tag{19}$$

With the above lemma, we are now able to analyze the static and dynamic convergence of adaptive blind equalizers.

IV. Static Convergence Analysis

If the equalizer is double-infinite, then at the global minimum of the cost function, the parameters of the equalizer

$$\{c_i\} = \{\pm \check{h}_{i-n_d}\}, \tag{20}$$

for some integer n_d. However, only FIR blind equalizer is used in practical systems. In this case, smart initialization strategies [4], [5], [9] will make the equalizer coefficients converge to a minimum $\{\tilde{c}_n : n = -N, \cdots, 0, \cdots, N\}$ of the cost function near the channel inverse such that $\tilde{y}_n - a_n$ is very small. Using first-order approximation to $\phi(.)$ at a_n, we can prove the following theorem.

Theorem 1: *If an FIR equalizer is used to equalize an FIR channel, then at the minimum near the channel inverse, the equalizer coefficient vector* $\tilde{\mathbf{c}} \triangleq [\tilde{c}_{-N}, \cdots, \tilde{c}_0, \cdots, \tilde{c}_N]^T$ *can be expressed as*

$$\tilde{\mathbf{c}} = f(0) R_f^{-1} \mathbf{h}. \tag{21}$$

where

$$\mathbf{h} = [0, \cdots, 0,\ h_M, \cdots, h_0, \cdots, h_{-M},\ 0, \cdots,\ 0]^T, \tag{22}$$

and

$$R_f \triangleq \mathcal{H}^T F \mathcal{H} \tag{23}$$

with

$$F = diag[\underbrace{f(1), \cdots, f(1)}_{M+N}, f(0), \underbrace{f(1), \cdots, f(1)}_{M+N}]. \tag{24}$$

and

$$f(0) \triangleq \frac{1}{\sigma^2} E\{\phi'(a_n) a_n^2\}, \quad f(1) \triangleq E\{\phi'(a_n)\}. \tag{25}$$

From the above theorem, the equalizer coefficient vector at the minimum of the cost function near the channel inverse is determined by (21).

For the channel with impulse response vector $\mathbf{h}$, the optimum equalizer (Wiener-Hopf filter) coefficient vector to minimize $E\{(y_n - a_n)^2\}$ is given by [7]

$$\mathbf{c}_o = R^{-1} \mathbf{h}, \tag{26}$$

where

$$R = \mathcal{H}^T \mathcal{H}. \tag{27}$$

Comparing (21) and (26), we have that the sufficient and necessary condition for $\tilde{\mathbf{c}} = \mathbf{c}_o$ for any FIR channel is

$$\mathcal{H}^T \mathcal{H} = \frac{1}{f(0)} \mathcal{H}^T F \mathcal{H}. \tag{28}$$

Since $\mathcal{H}$ is of full column rank for all no-zero $\mathbf{h}$, Equation (28) implies

$$f(0) = f(1), \tag{29}$$

which means

$$E\{\phi'(a_n) a_n^2\} = E\{\phi'(a_n)\} E\{a_n^2\}. \tag{30}$$

For Sato algorithm[14], decision-directed equalizers[10], [12], $\phi'(a_n) = 1$, and therefore, $\tilde{\mathbf{c}} = \mathbf{c}_o$. For Godard algorithm[6], $\phi(y) = y(y^2 - r)$ with $r = \frac{E\{a_n^4\}}{E\{a_n^2\}}$, therefore,

$$E\{\phi'(a_n) a_k^2\} = 3\sigma^4 - m_4, \tag{31}$$

and

$$E\{\phi'(a_n) a_n^2\} = 2m_4, \tag{32}$$

where

$$m_4 = E\{a_n^4\}. \tag{33}$$

Hence, if the channel input is binary, (30) is true and $\tilde{\mathbf{c}} = \mathbf{c}_o$. Otherwise, $\tilde{\mathbf{c}} \neq \mathbf{c}_o$.

The distortion due to the finite-length of equalizer is

$$\begin{aligned} D_f &\triangleq \| \tilde{\mathbf{s}} - \mathbf{e}_{M+N} \|^2 \\ &= \| \mathcal{H}\tilde{\mathbf{c}} - \mathbf{e}_{M+N} \|^2 . \end{aligned} \tag{34}$$

With the increase of the length of the blind equalizer, the global minimum of the cost function adopted by the equalization algorithm will be closer to the channel inverse. Hence, the distortion D_f will decrease.

V. Dynamic Convergence Analysis

When the blind equalization algorithms are implemented using stochastic gradient method, as are most blind equalizers, the blind equalizers will have an extra distortion $\epsilon_n \triangleq \hat{\mathbf{c}}^{(n)} - \tilde{\mathbf{c}}$ due to the gradient noise. Here, we study the stochastic dynamic convergence behavior of blind equalizers when the parameters of blind equalizers near the global minimum of the cost function. In our analysis, we will use the *independence assumption* which assumes that a_n and ϵ_n are statistically independent. Similar assumptions have also been used in the convergence analysis of LMS algorithm, decision-directed equalizer, and Sato algorithm. The references [7], [12], [20], [21] have given some good justification on the validation of this assumption.

By means of independence assumption, together with the first-order approximation, we are able to prove the following dynamic convergence theorem.

Theorem 2: *Let*

$$\tilde{R}_f \triangleq \mathcal{H}^T \tilde{F} \mathcal{H}, \tag{35}$$

with the largest eigenvalue λ_{max}*, and*

$$\tilde{R}_g \triangleq \mathcal{H}^T \tilde{G} \mathcal{H}, \tag{36}$$

with

$$\tilde{G} = \frac{1}{\sigma^2} E\{\mathbf{a}_n \phi^2(\tilde{y}_n) \mathbf{a}_n^T\}. \tag{37}$$

1) For any FIR blind equalization algorithm mean convergence behavior near the global minimum of the cost function satisfies

$$E\{\epsilon_n\} = (I - \mu\sigma^2 \tilde{R}_f)^n E\{\epsilon_0\}. \tag{38}$$

If the step-size μ *in iteration formula (5) or (7) satisfies*

$$0 < \mu < \frac{2}{\lambda_{max}\sigma^2}, \tag{39}$$

then

$$E\{\mathbf{c}^{(n)}\} \to \tilde{\mathbf{c}} \quad and \quad E\{\mathbf{s}^{(n)}\} \to \tilde{\mathbf{s}}, \tag{40}$$

2) The equalizer coefficient vector $\mathbf{c}^{(n)} \to \tilde{\mathbf{c}}$ *is not consistent and at the equilibrium near the minimum of the cost function, the correlation matrix* R_ϵ *of* ϵ *is uniquely determined by the following Lyapunov equation*

$$\tilde{R}_f R_\epsilon + R_\epsilon \tilde{R}_f = \mu \tilde{R}_g, \tag{41}$$

if $0 < \mu < \frac{1}{\lambda_{max}\sigma^2}$.

From the above theorem, the distortion of the equalized system due to gradient noise is

$$\begin{aligned} D_g &\triangleq E\{\| \mathbf{s} - \tilde{\mathbf{s}} \|^2\} \\ &= E\{\| \mathcal{H}\epsilon \|^2\} \\ &= \mathrm{tr}[\mathcal{H}^T R_\epsilon \mathcal{H}] \\ &= \mathrm{tr}[R R_\epsilon]. \end{aligned} \tag{42}$$

When an FIR equalizer is so long that $\{\tilde{c}_n \approx \check{h}_n\}$, $\{\tilde{y}_n \approx a_n\}$, then

$$\tilde{R}_f \approx R_f, \quad \tilde{R}_g \approx R_g, \tag{43}$$

where we have used the definitions

$$R_g \triangleq \mathcal{H}^T G \mathcal{H}, \tag{44}$$

and

$$g(0) \triangleq \frac{1}{\sigma^2} E\{\phi^2(a_n) a_n^2\}, \quad g(1) = E\{\phi^2(a_n)\}, \tag{45}$$

$$G \triangleq \mathrm{diag}[\underbrace{g(1), \cdots, g(1)}_{M+N}, g(0), \underbrace{g(1), \cdots, g(1)}_{M+N}]. \tag{46}$$

In this case, (41) becomes

$$R_f R_\epsilon + R_\epsilon R_f = \mu R_g. \tag{47}$$

For the blind equalization algorithms with $f(0) = f(1)$, $R_f = f(1)R$. Using (47), we have

$$D_g = \frac{1}{2f(1)} \mu \mathrm{tr}[R_g] \approx \frac{1}{2}\mu(2N+1)\frac{g(1)}{f(1)}. \tag{48}$$

For those blind equalizers with $f(0) \neq f(1)$, (48) can also be used to approximately estimate the average distortion introduced by gradient noise. According to (48), D_g is proportional to the step-size μ and the length of equalizer N. But, on the other hand, step-size affects the convergence speed of equalizers, i.e. the larger the μ, the faster it converges if μ is in the allowable range. Hence, when we select the step-size of an equalizer, we have to consider the trade-off between these two factors.

As we have seen, there are two sources of distortion. One is D_f in (4.24) due to the finite length of an equalizer, another is D_g in (48) due to the gradient noise. Once the step-size of a blind equalizer is set, there must be an optimum length that can be found for an FIR equalizer to minimize the total distortion $D = D_f + D_g$ since with the increase of the equalizer length, D_f decreases while D_g increases.

VI. Computer Simulations and Conclusion

Since approximation has been used in our theoretical analysis, we shall check the validity of our theory by computer simulations. Two computer simulation examples are presented in this section.

Example 1:

The channel input sequence $\{a_n\}$ is independent, uniformly distributed over $\{\pm a, \pm 3a\}$ ($a = 1/\sqrt{5}$ to make $E\{a_n^2\} = 1$). The impulse response of the channel is $h_n = 0.3^n u[n]$ with $u[n]$ being unit step function. An FIR equalizer with coefficients c_0 and c_1 is used to compensate for the channel distortion. The initial value of the equalizer coefficient vector is set to be

$$\mathbf{c}^{(0)} = [c_0^{(0)},\ c_1^{(0)}]^T = [1,\ 0]^T. \tag{49}$$

The Sato algorithm[14] is first used to adjust the coefficients of the equalizer. When the step-size $\mu = 0.002$, 10 trials of learning curves of $\mathbf{c}^{(n)}$ are shown in Figure 2. In this figure, the thick solid line is the theoretical average learning curve, the thick dot-dash lines are the theoretical one-standard-deviation lines. According to this figure, 10 trials of learning curves are almost within one standard deviation of the theoretical average learning curves for Sato algorithm.

Similar simulations have also been done for Godard algorithm[6]. The simulation results are shown in Figure 3, which also confirm our theoretical analysis.

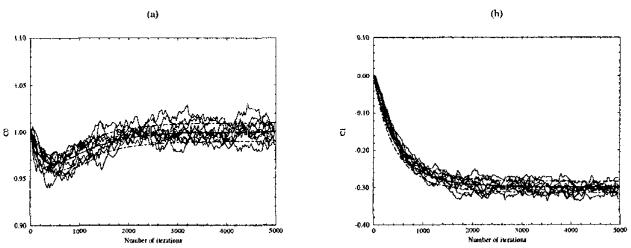

Fig. 2. 10 trials of learning curves of (a) c_0, and (b) c_1 for Sato algorithm using $\mu = 0.002$.

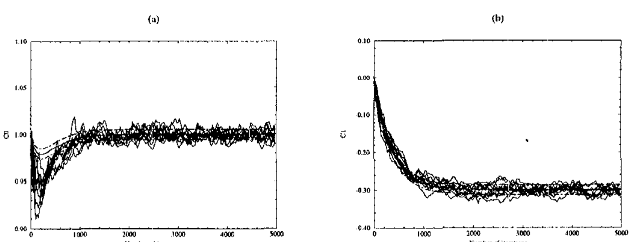

Fig. 3. 10 trials of learning curves of (a) c_0, and (b) c_1 for Godard algorithm using $\mu = 0.002$.

Example 2:

The channel input sequence in this example has the same statistical property as in Example 1. The channel impulse and frequency response are shown in Figure 4, which is a typical telephone channel [16]. The center-tap initialization strategy[5] is used for blind equalization algorithm.

When the Sato algorithm is used, the theoretical relationship between the total distortion and the length of equalizer for different step sizes are illustrated in Figure 5 (a), which indicates that the optimum length of Sato equalizer for this channel is between 15 and 25 dependent upon the step-size. Figure 5 (b) demonstrates the comparison between the theoretical results of $D_f + D_g$ and simulated results for step size $\mu = 0.002$.

The calculation and simulation results are given in Figure 6 for Godard algorithm. Because $g(1)/f(1)$ for Godard algorithm (0.169) is less than that for Sato algorithm (0.250) for 4-level PAM input, Godard algorithm should have less distortion than Sato should according to (48), which is confirmed by comparing Figure 5 and 6.

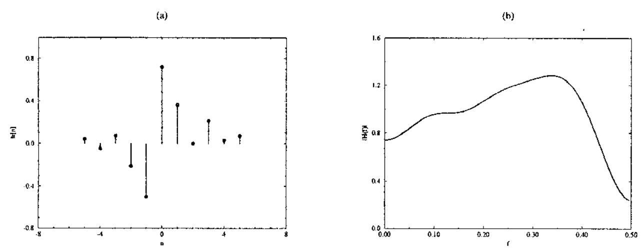

Fig. 4. (a) The impulse response, and (b) the frequency response of channel II.

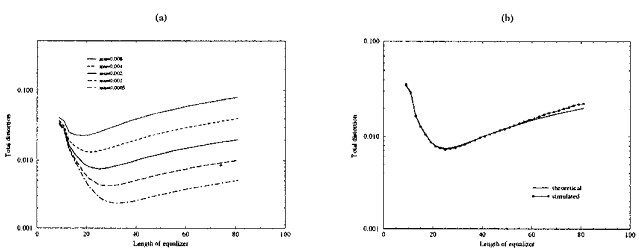

Fig. 5. Total distortion of equalized system (a) theoretical results for different step size μ, (b) simulation results for $\mu = 0.002$, using Sato algorithm.

We have studied the static and dynamic convergence behavior of adaptive blind equalizers in PAM digital communication systems based on the first-order approximation to the cost function of blind algorithms under the independence assumption. Most of the analysis results presented here can be extended to QAM digital communication systems. Our analysis result indicates that for a given channel and step-size, there is an optimal length for an equalizer to minimize the intersymbol interference. The results imply that a longer-length blind equalizer does not necessarily outperform a shorter one, as contrary to what is conventionally conjectured. The analysis results presented in this paper can be directly employed in the design of blind equalizer in practical communication systems.

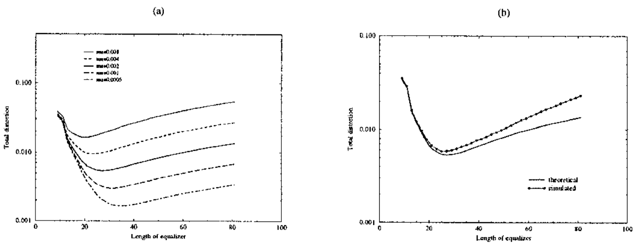

Fig. 6. Total distortion of equalized system (a) theoretical results for different step size μ, (b) simulation results for $\mu = 0.002$, using Godard algorithm.

References

[1] A. Benveniste, M. Goursat, and G. Ruget. "Robust identification of a nonminimum phase system: blind adjustment of a linear equalizer in data communications," *IEEE Trans. on Automatic Control*, AC-25:385–399, June 1980.

[2] A. Benveniste and M. Goursat, "Blind equalizers," *IEEE Transactions on Communications*, COM-32: 871-882, August, 1982.

[3] R. Cusani and A. Laurenti, "Convergence analysis of the CMA blind equalizer," *IEEE Transactions on Communications*, COM-43:1304-1307, Feb./March/April 1995.

[4] Z. Ding, R. A. Kennedy, B. D. O. Anderson and C. R. Johnson, Jr.. "Local convergence of the Sato blind equalizer and generalization under practical constrains," *IEEE Trans on Information Theory*, IT-39: 129–144, Jan. 1993.

[5] G. J. Foschini, "Equalization without altering or detecting data," *AT&T Technical Journal*, 64: 1885-1911, October, 1985.

[6] D. N. Godard, "Self-recovering equalization and carrier tracking in two-dimensional data communication systems," *IEEE Trans. on Communications*, COM-28, pp. 1867-1875, Nov. 1980.

[7] S. Haykin, *Adaptive Filter Theory*, Prentice-Hall, 1991

[8] R. A. Kennedy, B. O. Anderson and R. R. Bitmead, "Stochastic Dynamics of Blind Decision Feedback Equalizer Adaptation," *Adaptive Systems in Control and Signal Processing 1989* pp.579-584, IFAC Symposia Series, 1990.

[9] Y. Li and Z. Ding, "Convergence analysis of finite length blind adaptive equalizer," *IEEE Trans on Signal Processing*, SP-42:2120-2129, September 1995.

[10] O. Macchi and E. Eweda, "Convergence analysis of self-adaptive equalizers," *IEEE Transactions on Information Theory*, IT-30:161-176, March 1984.

[11] V. J. Mathews and H. Cho, "Improved convergence analysis off stochastic gradient adaptive filters using the sign algorithm," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, ASSP-35: 450-454, April, 1987.

[12] J. E. Mazo, "Analysis of Decision-Directed Equalizer Convergence," *The Bell System Technical Journal*, Vol.59, No.10, pp1857-1876, December 1980.

[13] Y. Sato, "A method of self-recovering equalization for multi-level amplitude modulation," *IEEE Trans. on Comm.* COM-23: 679-682, June 1975

[14] C. K. Chan and J. J. Shynk, "Stationary Points of the Constant Modulus Algorithm for Real Gaussian Signals," *IEEE Transactions on Signal Processing*, ASSP-38:2176-2180, 1990.

[15] J. Proakis, *Digital Communications* (2nd Edition), McGraw-Hill 1989.

[16] J. R. Treichler and B.G. Agee, "A new approach to multipath correction of constant modulus signals," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, ASSP-31: 349-372, April, 1983.

[17] S. Verdu, B. D. O. Anderson, R. A. Kennedy, "Blind equalization without gain identification," *IEEE Transactions on Information Theory*, IT-39:292-297, Jan. 1993.

[18] V. Weerackody, S. A. Kassam, and K. R. Laker, "Convergence analysis of an algorithm for blind equalization," *IEEE Transactions on Communications*, Vol.39:856-865, June, 1991.

[19] B. Windrow *et al*, "Stationary and Nonstationary Learning Characteristics of the LMS Adaptive Filter," *Proceedings of the IEEE*, Vol.64:1151-1162, August 1976.