

# Length- and Cost-Dependent Local Minima of Unconstrained Blind Channel Equalizers

Ye Li, K. J. Ray Liu, and Zhi Ding

**Abstract**—Baud-rate linear blind equalizers may converge to undesirable stable equilibria due to different mechanisms. One such mechanism is the use of linear FIR filters as equalizers. In this paper, it is shown that this type of local minima exist for all unconstrained blind equalizers whose cost functions satisfy two general conditions. The local minima generated by this mechanism are thus called *length-dependent local minima*. Another mechanism is generated by the cost function adopted by the blind algorithm itself. This type of local minima are called *cost-dependent local minima*. It shall be shown that several well-designed algorithms do not have cost-dependent local minimum, whereas other algorithms, such as the decision-directed equalizer and the stop-and-go algorithm (SGA), do. Unlike many existing convergence analysis, the convergence of the Godard algorithms (GA's) and standard cumulant algorithms (SCA's) under Gaussian noise is also presented here. Computer simulations are used to verify the analytical results.

## I. INTRODUCTION

**B**LIND channel equalization is a useful tool for the removal of intersymbol interference (ISI) in digital communication systems when training sequences are costly or impractical. Blind (adaptive) channel equalization algorithms are usually designed to minimize cost functions based on statistics of channel output signals. If the underlying cost function of an algorithm has local minima in addition to the global one(s), undesirable local convergence becomes possible.

As discovered in [4], local convergence of baud-rate blind equalization algorithms may be the result of two different causes. One cause is the standard use of finite length equalizer filter. As will be shown in this paper, this kind of local minima exist for all finite length baud-rate blind equalizers without filter parameter constraint. They are thus called *length-dependent local minima*. Another kind of local minima result from poor selections of cost functions. They can exist even under the ideal “doubly infinite” [4] equalizer filters. Local minima generated by this mechanism are called *cost-dependent*

*local minima*. The latter kind of local minima do not exist for well-designed cost functions.

Due to the recent upsurge of interest in blind equalization, many different adaptive algorithms have been proposed. The Sato algorithm (SA) [18] was the first known blind equalization algorithm. It was later generalized by Benveniste, Goursat, and Ruget [1] into a set of what we call “BGR algorithms” (BGRA). It has been proved in [5] that both length and cost-dependent local minima exist for SA and the BGRA [5]. The local convergence of the decision directed equalizer (DDE) and the stop-and-go algorithm (SGA) [17] has been analyzed in [11], [14], and [15]. Since the DDE and SGA are identical to the SA for channels with binary inputs, their cost-dependent local convergence is also given by [5].

On the other hand, the Godard algorithms (GA's) are a different generalization of the SA by Godard [10]. The Godard algorithm was later extended into a set of Shalvi–Weinstein algorithms (SWA's) [25], [26]. Both the GA and SWA are well-known and effective algorithms. The analysis of GA is given in [3], [4], and [7]. It is shown in [3] that Godard cost function also has length-dependent local minima (LDLM), but it has no cost-dependent local minima (CDLM) [7]. Based on results in [12] that a one-to-one correspondence exists between the minima of the GA and SWA, the SWA has the same convergence performance as the GA.

Although some convex cost functions can be designed for adaptive blind equalizers under specific parameter constraints [23], [24], [27] to avoid local convergence, these algorithms tend to be rather slow due to their  $l_\infty$  nature of the cost functions. For equalizer filters without constraints, convergence analyzes of [3]–[5], [12], [14], and [15] have shown LDLM to exist for many known algorithms such as BGRA, SA, GA, SWA, SGA, and DDE. Therefore, a natural question is whether or not it is possible to design good cost functions so that LDLM do not exist even for the common unconstrained FIR equalizer filters. Moreover, all the above convergence analyzes of blind algorithms are based on the noiseless channel assumption for analytical simplicity. There are no analytical results under noise yet. In fact, there have been some conjecture that channel noises may help equalizer parameters to escape some shallow local minima. Since channel noise is present in all practical communication systems even it is often small, convergence analysis of blind equalization algorithms must be carried out under channel noise. In this paper, we address these two important questions. We will study whether LDLM can be eliminated by general blind equalizers under no parameter constraints.

Manuscript received April 24, 1995; revised April 30, 1996. This work was supported, in part, by NSF grants MIP9309506, MIP9457397, and MIP-9210100 and the US Army Research Office. The associate editor coordinating the review of this paper and approving it for publication was Prof. Jose Carlos M. Bermudez.

Y. Li was with the Department of Electrical Engineering, University of Maryland, College Park, MD 20742 USA. He is now with the Wireless Systems Research Department, AT&T Laboratories Research, Holmdel, NJ 07733-0400 USA.

K. J. R. Liu is with the Department of Electrical Engineering, University of Maryland, College Park, MD 20742 USA.

Z. Ding is with the Department of Electrical Engineering, Auburn University, Auburn, AL 36849-5201 USA.

Publisher Item Identifier S 1053-587X(96)08239-6.

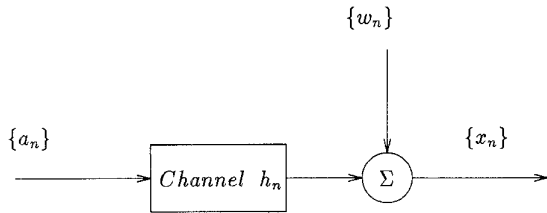


Fig. 1. Baseband representation of PAM communication system.

We will also study the equalizer convergence under channel noise.

This paper is presented as following. In Section II, we briefly introduce the pulse amplitude modulation (PAM) blind equalization system. In Section III, we show that LDLM will be present for blind equalization algorithms whose cost functions satisfy two general conditions without equalizer parameter constraints. In Section IV, we show that the standard cumulant algorithm (SCA) has no CDLM. We then study the global convergence of GA and SCA under white Gaussian noise and derive the mean square error (MSE) for the system output at the desired global convergence. Next, we demonstrate the existence of CDLM for DDE and SGA under nonbinary channel input signals in Section V. Our analysis results are confirmed by computer simulations in Section VI.

## II. BLIND EQUALIZATION

### A. Equalization in Digital Communication Systems

Without loss of generality, we consider a baud-rate sampled baseband representation of the pulse-amplitude-modulation (PAM) communication system as shown in Fig. 1. A sequence of independent, identically distributed (i.i.d.) digital signal  $\{a_n \in \mathcal{R}\}$  is sent through a channel  $h_n$  exhibiting linear distortion. The resulting output signal  $x_n$  can be expressed as

$$x_n = \sum_{k=-\infty}^{+\infty} a_{n-k} h_k + w_n \quad (2.1)$$

where  $w_n$  is additive channel noise. We will assume that the channel

$$H(\omega) \triangleq \sum_n h_n e^{-j\omega n}$$

is bounded-input-bounded-output (BIBO) stable, which implies that  $\sum_n |h_n| < \infty$ .

A linear channel equalizer

$$\Theta(\omega) \triangleq \sum_n \theta_n e^{-j\omega n}$$

is applied to the channel output  $\{x_n\}$  in order to eliminate the ISI. For the equalizer to be BIBO and practically implementable, it is necessary that  $\sum_n |\theta_n| < \infty$ . The equalizer parameters  $\{\theta_n\}$  are subject to adaptation via some algorithm to be determined. If we let

$$S(\omega) \triangleq H(\omega)\Theta(\omega) = \sum_n s_n e^{-j\omega n} \quad (2.2)$$

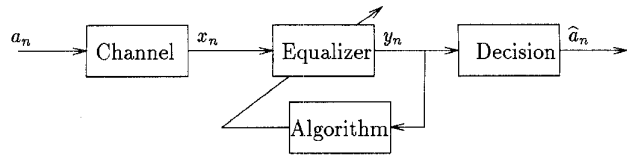


Fig. 2. Diagram of typical blind equalization system.

then

$$s_n \triangleq \sum_k h_k \theta_{n-k}. \quad (2.3)$$

The overall channel-equalizer system is obviously also BIBO  $\sum_n |s_n| < \infty$ .

From Fig. 2, the equalizer output can be written as

$$\begin{aligned} y_n &= \sum_{k=-\infty}^{\infty} \theta_k x_{n-k} \\ &= \sum_{k=-\infty}^{\infty} s_k a_{n-k}. \end{aligned} \quad (2.4)$$

In the above expression, we have ignored the channel noise. Our analysis on noise effect will be presented in Section IV-B.

In blind equalization, the original sequence is unknown to the receiver except for its probabilistic or statistical properties over the known alphabet  $\mathcal{A}$ , which is real for PAM and complex for quadrature amplitude modulation (QAM). Usually, this signal constellation  $\mathcal{A}$  is symmetric, resulting in symmetric statistics for the i.i.d. input data. Thus, the recovered data from blind equalization will be similarly subjected to a sign ambiguity, and the best possible result is

$$H(\omega)\Theta(\omega) = \pm e^{-jn_d\omega}, \quad (2.5)$$

for some integer  $n_d$ .

### B. Adaptive Blind Equalization Algorithms

Adaptive blind equalization algorithms are typically designed to minimize well-defined statistical cost functions through unconstrained optimization of equalizer parameters  $\theta_n$ . Cost functions are based on statistics of equalizer output  $y_n$ , which can also be viewed as functions of  $\{\dots, s_{-1}, s_0, s_1, \dots\}$ . As stated in (2.5), the cost functions should be designed to have global minimum points at

$$\{s_n\} = \pm\{\delta[n - n_d]\} \quad \text{for all } n_d = 0, \pm 1, \pm 2, \dots \quad (2.6)$$

to attain perfect equalization.

The GA [10], which is also known as the constant modulus algorithm (CMA), is one such widely used algorithm. The algorithm adjusts the equalizer parameter  $\{\theta_n\}$  by minimizing the Godard cost function defined as

$$f_{GA}(\theta) \triangleq \frac{1}{4} E\{(|y_n|^2 - \gamma)^2\}, \quad \gamma = \frac{E\{|a_n|^4\}}{E\{|a_n|^2\}}. \quad (2.7)$$

Let  $C_{y_n}^p$  be the  $p$ th-order cumulant of  $y_n$  defined as

$$C_y^p \triangleq \left(-j \frac{d}{dt}\right)^p \log E\{e^{jty}\}|_{t=0}. \quad (2.8)$$

The SCA's [6] are defined by minimizing the cost function

$$f_{SCA}(\theta) \triangleq -|C_{y_n}^{2m}| \text{ subject to } |C_{y_n}^2| = |C_{a_n}^2| \text{ where } m > 1. \quad (2.9)$$

As shown in [26], the SWA [25] is a special case ( $m = 2$ ) of the SCA.

Adaptive algorithms are typically obtained as stochastic gradient descent method to minimize the cost function. Given a cost function that can be written as

$$f(\theta) = E\{\Psi(y_n)\}, \text{ in which } \frac{\partial \Psi(x)}{\partial x} = \psi(x) \quad (2.10)$$

an on-line adaptive equalization algorithm adjusts the parameters of the equalizer via

$$\hat{\theta}_k^{(n+1)} = \hat{\theta}_k^{(n)} - \mu \psi(y_n) x_{n-k}, \quad k = 0, \pm 1, \pm 2, \dots \quad (2.11)$$

Here,  $\mu$  is a small adaptation step size, and  $\psi(\cdot)$  is sometimes referred to as the *error function*.

Some blind equalization algorithms are directly defined from (2.11). Let

$$\psi(y_n) = w(y_n)(y_n - \hat{y}_n). \quad (2.12)$$

If  $w(y_n) = 1$  for all  $y_n$  in (2.12), then (2.11) is the DDE [15]. If  $w(y_n)$  is defined as

$$w(y_n) = \begin{cases} 1 & \text{if } \text{sign}(y_n - \hat{y}_n) = \text{sign}(y_n - \beta \text{sign}(y_n)) \\ 0 & \text{otherwise} \end{cases} \text{ for some } \beta > 0 \quad (2.13)$$

the algorithm defined by (2.11)–(2.13) is the SGA [17].

In addition to the above algorithms based on nonconvex cost functions, some convex cost functions have also been presented for blind equalization that necessarily require parameter constraint(s) to avoid trivial solutions [23], [24], [27]. Although constrained convex cost minimization avoid the problem of local minima, they are far from ideal in terms of bias and the speed of convergence. Because their convergence behaviors are well regulated due to convexity, we only deal with the unconstrained blind equalization algorithms in this paper.

### III. LENGTH-DEPENDENT LOCAL MINIMA IN BLIND EQUALIZATION

Since finite parameterization is a practical necessity for equalizers, the resulting cost function under finite parameterization may have undesirable minima. Undesirable local convergence behavior has been shown for the GA (CMA case) [3], [4], the BGRA [5], and the SWA [12]. In this section, we shall show that all unconstrained blind equalization algorithms will have local minima if the equalizer is implemented as an FIR filter.

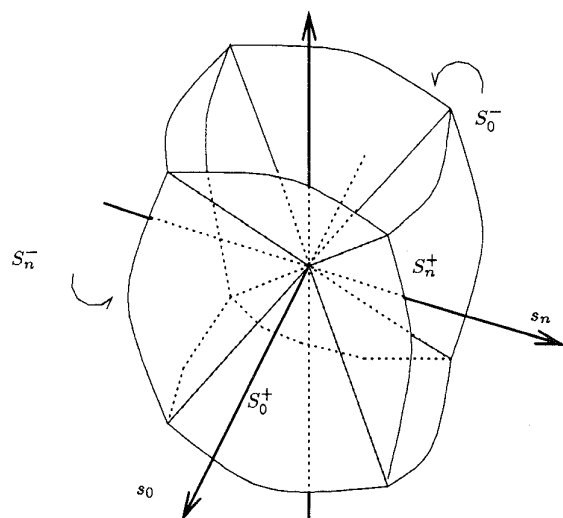


Fig. 3. Geometric structure of unique global minimum cones.

#### A. Useful Definitions

We first introduce several useful notations and definitions. Let  $\ell^1(\mathcal{R})$  denote the set of all real sequences  $\mathbf{u} = \{\dots, u_{-1}, u_0, u_1, \dots\}$  with finite  $\ell^1$ -norm, i.e.,

$$\|\mathbf{u}\| = \sum_{n=-\infty}^{\infty} |u_n| < \infty. \quad (3.1)$$

The BIBO assumption implies that

$$\mathbf{h} \triangleq (\dots, h_{-1}, h_0, h_1, \dots) \in \ell^1(\mathcal{R}), \quad (3.2)$$

$$\boldsymbol{\theta} \triangleq (\dots, \theta_{-1}, \theta_0, \theta_1, \dots) \in \ell^1(\mathcal{R}) \quad (3.3)$$

and

$$\mathbf{s} \triangleq (\dots, s_{-1}, s_0, s_1, \dots) \in \ell^1(\mathcal{R}). \quad (3.4)$$

The cost function can be expressed in terms of  $\mathbf{s}$  or  $\boldsymbol{\theta}$  as a functional on  $\ell^1(\mathcal{R})$ .

The *hyperball*  $\Phi(\mathbf{o}, \rho)$  with center  $\mathbf{o}$  and radius  $\rho$  is defined as

$$\Phi(\mathbf{o}, \rho) = \{\mathbf{s} \in \ell^1(\mathcal{R}): \|\mathbf{s} - \mathbf{o}\| \leq \rho\}. \quad (3.5)$$

The *unique global minimum cones* are defined as

$$\begin{aligned} S_n^+ &\triangleq \{\mathbf{s} \in \ell^1(\mathcal{R}): s_n > 0 \text{ and } s_n > |s_k| \text{ for all } k \neq n\} \\ S_n^- &\triangleq \{\mathbf{s}: -\mathbf{s} \in S_n^+\}. \end{aligned} \quad (3.6)$$

The geometric structure of unique global minimum cones is shown in Fig. 3.

Their boundaries are, respectively, given by

$$\begin{aligned} B_n^+ &\triangleq \{\mathbf{s} \in \ell^1(\mathcal{R}): s_n > 0 \text{ and } s_n \geq |s_k| \\ &\quad \text{and the equality holds for some } k \neq n\} \\ B_n^- &\triangleq \{\mathbf{s}: -\mathbf{s} \in B_n^+\}. \end{aligned} \quad (3.7)$$

In each cone  $S_n^+$  (or  $S_n^-$ ), the cost function  $f(\mathbf{s})$  has only one global minimum point  $\mathbf{e}_n^+$  (or  $\mathbf{e}_n^- = -\mathbf{e}_n^+$ ), where

$$\mathbf{e}_n^+ \triangleq (\dots, 0, \underbrace{1}_{(n\text{-th})}, 0, \dots). \quad (3.8)$$

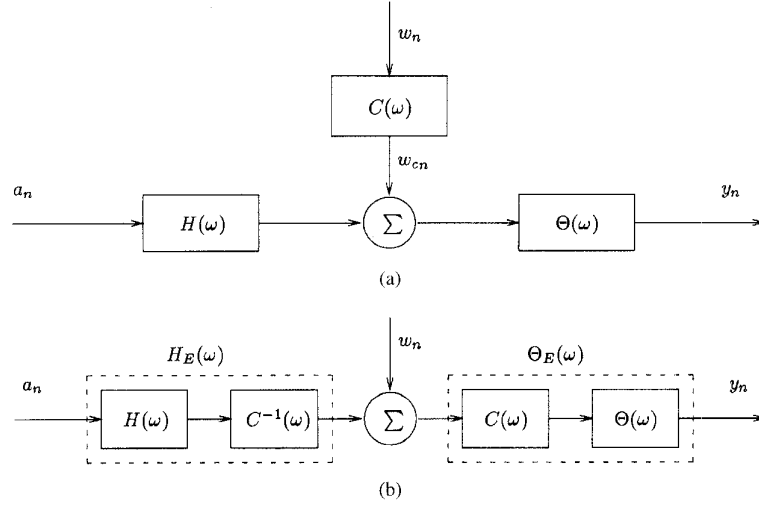


Fig. 4. (a) System with colored channel noise and (b) its equivalent white noise model.

### B. Local Convergence Characterization

With the above definitions, we now proceed to characterize the local convergence of blind equalization algorithms without parameter constraint. For reasonably designed cost functions, we will require that they satisfy the following two conditions:

- C1:  $e_n^+$  and  $e_n^-$  for  $n = \dots, -1, 0, 1, \dots$  are the only global minimum points of the cost function  $f(\mathbf{s})$ .
- C2: The cost function  $f(\mathbf{s})$  is continuous on  $\ell^1(\mathcal{R})$ .

First, we denote the minimum on the boundary

$$f_{mb} \triangleq \min_{\mathbf{s} \in B_n^+} f(\mathbf{s}). \quad (3.9)$$

Then, based on C1, it is necessary that

$$f_{mb} > f(e_n^+). \quad (3.10)$$

In addition, from C2,

$$O_n^+ \triangleq \{\mathbf{s} \in S_n^+ : f(\mathbf{s}) < f_{mb}\} \quad \text{for } n = \dots, -1, 0, 1, \dots \quad (3.11)$$

is an open set in  $\ell^1(\mathcal{R})$  containing  $e_n^+$  since  $f(\mathbf{s})$  is continuous. Hence, there exists a  $\rho > 0$  such that

$$\Phi(e_n^+, \rho) \in O_n^+. \quad (3.12)$$

It is apparent from the definition that the maximum cost in the hyperball  $\Phi(e_n^+, \rho)$  is lower than the cost on the cone boundary

$$\max_{\mathbf{s} \in \Phi(e_n^+, \rho)} f(\mathbf{s}) < f_{mb}. \quad (3.13)$$

In most linear equalization systems, a finite-length equalizer filter

$$\sum_{n=N_1}^{N_2} \theta_n z^{-n}$$

with coefficients  $\theta_n$  is normally used as an equalizer in practical communication systems, which means only these  $N_2 - N_1 + 1$  coefficients can be adjusted, and the rest of

them are fixed to be zero. Hence, not all  $\mathbf{s} \in \ell^1(\mathcal{R})$  can be attained by the equalizer, and the *attainable set*  $A$  is

$$A \triangleq \{\mathbf{s} \in \ell^1(\mathcal{R}) : s_n = \sum_{k=N_1}^{N_2} \theta_k h_{n-k}, \theta_k \in \mathcal{R}\}. \quad (3.14)$$

Since FIR equalizer restricts the algorithm minimization of  $f(\mathbf{s})$  to  $A$ , local minimum points can be thus generated even if  $f(\mathbf{s})$  only has desired global minima without restriction. We now give a proof here for cost functions satisfying the two conditions.

Consider a PAM channel impulse response

$$h_n = \alpha^n u[n], \quad \text{where } \alpha = \frac{\rho}{1+\rho} \quad \rho > 0 \quad (3.15)$$

where  $u[n]$  is the unit step function. It is a stable first-order autoregressive system that would only require a first-order moving-average system to perfectly equalize the channel. If an  $N$ -tap filter with coefficients  $\theta_0, \theta_1, \dots, \theta_{N-1}$  is used as equalizer, the attainable set can be expressed as

$$A = \{\mathbf{s} : s_n = \sum_{k=0}^n \theta_k \alpha^{n-k}, \theta_k \in \mathcal{R}\}. \quad (3.16)$$

It is easy to check that  $e_n^+ \in A$  for  $n = 0, \dots, N-2$  and  $e_{N-1}^+ \notin A$ . However,  $A \cap \Phi(e_{N-1}^+, \rho)$  is not empty since if  $\theta_0 = \theta_1 = \dots = \theta_{N-2} = 0, \theta_{N-1} = 1$ , then  $\{\mathbf{s}_n\} \in A$  and  $\|\mathbf{s} - e_{N-1}^+\| = \rho$ . Therefore, if  $\theta_0 = \theta_1 = \dots = \theta_{N-2} = 0, \theta_{N-1} = 1$  is used as the initial setting of the equalizer, the steepest descent line (s.d.l) will never across the boundary of the cone  $S_{N-1}^+$  because of (3.13). Hence, the equalizer with this initial setting and small stepsize will converge to some minimum in  $S_{N-1}^+$ , which is obviously not the global minimum  $e_{N-1}^+$  in this cone because it is not in  $A \cap S_{N-1}^+$ . Hence, we have proved the following theorem.

**Theorem 3.1:** If the cost function of a blind equalizer satisfies conditions C1 and C2, the undesirable stable minima exist for the finite length equalizer.

*Remarks:*

- 1) In the proof of the theorem, we did not restrict the distribution of the channel input set. Hence, the theorem can be used for any finite length equalizer with input symbols of any non-Gaussian distribution as long as the cost function satisfies conditions *C1* and *C2*. Clearly, condition *C1* is not satisfied for Gaussian input for which baud-rate blind equalization is impossible.
- 2) The above theorem is only concerned with the existence of local minima. If the equalizer is initialized to have  $\theta_0 = 1, \theta_1 = 0$ , then the equalizer will converge to some minimum point inside  $\Phi(e_0^+, \rho)$ . However, the analysis does not specify whether  $e_0^+$  is a unique minimum inside  $\Phi(e_0^+, \rho)$  for the cost function.
- 3) Since conditions *C1* and *C2* must be satisfied by all reasonably designed cost functions without constraints, Theorem 3.1 shows that these unconstrained blind equalization algorithms must possess local minima caused by the finite filter length. Such local minima are thus called *length-dependent local minima*. When the equalizer parameters are constrained, however, condition *C1* is usually not satisfied. Hence, Theorem 3.1 is no longer applicable to such algorithms as developed in [23], [24], [27].

## IV. CONVERGENCE ANALYSIS OF THE SCA AND THE GA

The convergence analysis of the GA and SWA has been presented in [3], [4], and [12] for noiseless channels. In this section, these analytical results are first extended to the SCA. Moreover, the global convergence of SCA and the GA under white Gaussian channel noise will also be studied (see Fig. 4).

## A. Convergence of SCA without Channel Noise

First, we define the normalized kurtosis of a random process  $x_n$  as

$$K_{x_n}^{2m} = \frac{C_{x_n}^{2m}}{(C_{x_n}^2)^m}. \quad (4.1)$$

Similar to the Godard equalizer and the Shalvi-Weinstein equalizer [12], the finite-length equalizer using the SCA have the following convergence properties.

*Theorem 4.1:* For (finite-length or infinite-length) standardized cumulant equalizers, suppose the initial equalizer parameter setting causes the initial overall parameter vector to satisfy  $s^{in} \in A \cap S_n^\pm$ . If the initial equalizer output  $y_n$  satisfies

$$\frac{|K_{y_n}^{2m}|}{|K_{a_n}^{2m}|} > 0.5, \quad (4.2)$$

then under very small minimization step size, the equalizer will cause  $s$  to converge to a minimum point inside  $A \cap S_n^\pm$ .

*Theorem 4.2:* Let  $A$  be the attainable set of a finite-length standardized cumulant equalizer. Then, we have the following:

- 1) If  $e_n^\pm \in A \cap S_n^\pm$ , then there is only one minimum point  $e_n^\pm$  in  $A \cap S_n^\pm$ , and there is no minimum point on the boundary  $A \cap B_n^\pm$ .

- 2) If  $e_n^\pm$  is near but not in  $A \cap S_n^\pm$  (not attainable), then there must exist only one minimum point in  $A \cap S_n^\pm$  near  $e_n^\pm$  while all other possible minima are near the boundary  $A \cap B_n^\pm$ .

The proofs of the above two theorems are similar to the proofs of Theorem 5.1 and Theorem 5.2 in [12], respectively, and are thus omitted here. The special case of Theorem 4.2 (SCA with  $m = 2$ ) is analyzed in [25]. The similar analysis for BGRA under a sub-Gaussian input signal has been given in [1].

Based on the above theorems, SCA also has length-dependent local minima resulting from the finite FIR equalizer length. With these theorems, the initialization strategy for GA and SWA discussed in [12] can also be used for SCA. If the length of an SCA equalizer is double-infinite (an impractical abstraction), then  $e_n^\pm$  for all  $n$  are in the attainable set  $A$ . From Theorem 4.2, the SCA equalizer will converge to some  $e_n^+$  or  $e_n^-$ . Hence, SCA has no cost-dependent local minima.

## B. Global Minima of the GA and SCA under Gaussian Channel Noise

In this part of the paper, we study the biasing effect of Gaussian channel noise on the global minima of the GA and SCA. We will first study the convergence of the Godard algorithm under white Gaussian noise. To facilitate the analysis, we will assume that the signal-to-noise ratio is very high.

When there is channel noise, the equalizer output can be expressed as

$$y_n = \sum_{k=-\infty}^{\infty} s_k a_{n-k} + \sum_{k=-\infty}^{\infty} \theta_k w_{n-k}. \quad (4.3)$$

Hence, from the definition of (2.7), the Godard cost function can be written as

$$\begin{aligned} g(\boldsymbol{\theta}) &= f(\mathbf{s}) \\ &= \frac{1}{4} E\{(|y_n|^2 - \gamma)^2\} \\ &= \frac{1}{4} m_2^2 \left\{ -(3-r) \sum_n s_n^4 + 3 \left( \sum_n s_n^2 \right)^2 \right. \\ &\quad - 2r \sum_n s_n^2 + r^2 + 6\eta \sum_n s_n^2 \sum_n \theta_n^2 \\ &\quad \left. + 3\eta^2 \left( \sum_n \theta_n^2 \right)^2 - 2r\eta \sum_n \theta_n^2 \right\} \end{aligned} \quad (4.4)$$

where

$$r = \frac{m_4}{m_2^2} \quad \text{and} \quad \eta = \frac{\sigma^2}{m_2}. \quad (4.5)$$

It can be verified from the derivative of the above equation that the global minima of the Godard equalizer will no longer be  $e_n^\pm$ . On the other hand, if the signal-to-noise ratio is high (small  $\eta$ ), then the global minima will only be slightly perturbed away from  $e_n^\pm$ .

We denote the perturbed global minima as  $e_n^\pm + \Delta s$ . The perturbed minima in terms of equalizer parameters can be

denoted as  $\tilde{\mathbf{h}} + \Delta\boldsymbol{\theta}$ , where  $\{\tilde{h}_n\} = \mathcal{Z}^{-1}\{H^{-1}(z)\}$ , and  $\Delta s_n = \sum_k h_k \Delta\theta_{n-k}$ . Using Taylor expansion, we have

$$\begin{aligned} f(\mathbf{s}) = & f(\mathbf{e}_0^+) + \frac{1}{2}m_2^2 \left[ (3-r) \sum_n (\Delta s_n)^2 \right. \\ & \left. + 3(r-1)(\Delta s_0)^2 + O(\|\Delta\mathbf{s}\|^2) \right] \\ & + m_2^2 \eta \left[ (3-r) \sum_n t_n \Delta\theta_n + 3d_2 \sum_n h_{-n} \Delta\theta_n \right. \\ & \left. + O(\|\Delta\boldsymbol{\theta}\|) \right] \end{aligned} \quad (4.6)$$

where

$$d_m = \frac{1}{2\pi} \int_{-\pi}^{\pi} \frac{1}{|H(\omega)|^m} d\omega. \quad (4.7)$$

Ignoring the  $O(\|\Delta\mathbf{s}\|^2)$  and  $\eta O(\|\Delta\boldsymbol{\theta}\|)$  terms in the above expression, a direct calculation yields that the minimum of the cost function will satisfy

$$\begin{aligned} (3-r)\Delta S(\omega)H^*(\omega) + 3(r-1)\Delta s_0 H^*(\omega) \\ + (3-r)\eta H^{-1}(\omega) + 3\eta d_2 H^*(\omega) \approx 0 \end{aligned} \quad (4.8)$$

where  $\Delta S(\omega)$  is the Fourier transform of  $\{\Delta s_n\}$ , and  $*$  denotes complex conjugate. Hence,

$$\Delta S(\omega) \approx - \left( \frac{1}{|H(\omega)|^2} + \frac{3(2-r)}{2r} d_2 \right) \eta. \quad (4.9)$$

The transfer functions of the equalized system and the equalizer are

$$S(\omega) \approx 1 - \left( \frac{1}{|H(\omega)|^2} + \frac{3(2-r)}{2r} d_2 \right) \eta \quad (4.10)$$

and

$$\Theta(\omega) \approx \frac{S(\omega)}{H(\omega)}. \quad (4.11)$$

The MSE after equalization will be

$$(\text{MSE})_{GA} \approx d_2 \eta - \left[ d_4 - \frac{9(2-r)^2}{4r^2} d_2^2 \right] \eta^2. \quad (4.12)$$

Similar derivation for SCA yields the transfer functions of the overall system and the equalizer as

$$S(\omega) \approx 1 - \left[ \frac{1}{|H(\omega)|^2} - \frac{1}{2} d_2 \right] \eta \quad (4.13)$$

and

$$\Theta(\omega) \approx \frac{S(\omega)}{H(\omega)} \quad (4.14)$$

respectively. The resulting MSE for SCA is

$$(\text{MSE})_{SCA} \approx d_2 \eta - (d_4 - \frac{1}{4} d_2^2) \eta^2. \quad (4.15)$$

As a comparison, we also determine the system MSE when the channel inverse and the Wiener filter are used as equalizers. The resulting MSE's are

$$(\text{MSE})_{\text{inv}} = d_2 \eta, \quad (4.16)$$

$$(\text{MSE})_{\text{Win}} \approx d_2 \eta - d_4 \eta^2 \quad (4.17)$$

respectively. From (4.15), the MSE of SCA will be between the MSE of (optimum) Wiener filter and channel inverse, i.e.,

$$(\text{MSE})_{\text{inv}} > (\text{MSE})_{SCA} > (\text{MSE})_{\text{Win}}. \quad (4.18)$$

In addition, , if the channel input PAM signal has over two levels, then  $1.68 \leq r \leq 1.8$ . Consequently,

$$0 < d_4 - \frac{9(2-r)^2}{4r^2} d_2^2 < d_4 \quad (4.19)$$

and

$$(\text{MSE})_{\text{inv}} > (\text{MSE})_{GA} > (\text{MSE})_{\text{Win}}. \quad (4.20)$$

### C. Global Convergence Under Colored Gaussian Noise

When the channel noise  $w_{cn}$  is a regular, colored Gaussian process with power spectrum  $W(\omega)$ , there exists a minimum-phase function  $C(\omega)$  such that

$$W(\omega) = C(\omega)C^*(\omega). \quad (4.21)$$

Therefore, the colored Gaussian noise  $w_{cn}$  is equivalent to the output of a system with transfer function  $C(\omega)$  driven by white Gaussian input  $w_n$ . The system with regular colored channel noise can be demonstrated by Fig. 7(a), which is equivalent to Fig. 7(b). In Fig. 7(b),

$$H_E(\omega) = \frac{H(\omega)}{C(\omega)} \quad (4.22)$$

and

$$\Theta_E(\omega) = \Theta(\omega)C(\omega). \quad (4.23)$$

Hence, the results for white Gaussian noise can be directly applied.

## V. COST-DEPENDENT LOCAL MINIMA OF DDE AND SGA

For blind equalization algorithms involving nondifferentiable cost functions, the convergence analysis becomes more difficult since a general analytical expression of the cost function cannot be obtained. The local convergence of the DDE implemented with an FIR filter is demonstrated in [11], [14], and [15], which reveals the length-dependent local convergence of DDE. In this section, we will illustrate the cost-dependent local convergence of the DDE and the SGA by specifying their cost functions in some operating regions.

For cost-dependent local convergence, consider a noiseless PAM communication system with a double-infinite-length equalizer for which the entire  $s$  space is the attainable set [5]  $A = \ell^1(\mathcal{R})$ . Consider an i.i.d. PAM channel input uniformly distributed over the set  $\{-3, -1, 1, 3\}$ . If a doubly infinite length SGA equalizer with  $\beta = 2$  is used in the system, we will show that  $\{s_n\} = \{\frac{13}{31}(\delta[n-1] + \delta[n] + \delta[n+1])\}$  is one of the local minimum of the SGA.

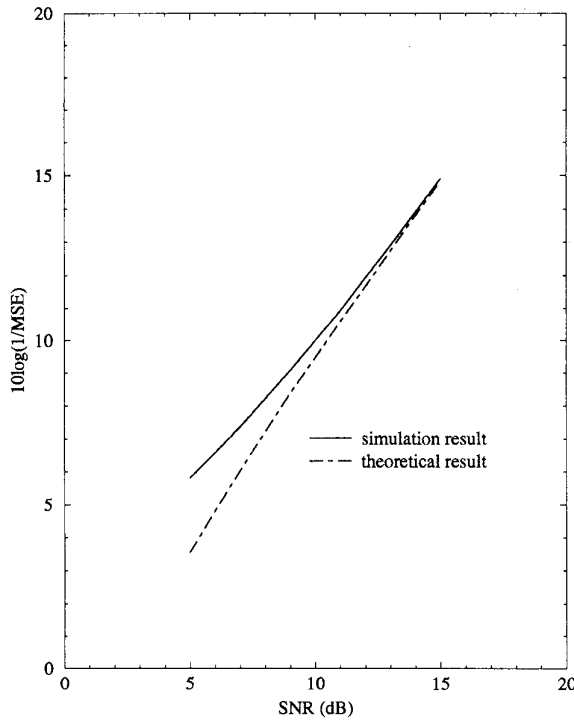


Fig. 5. MSE of the equalizer output when the channel input is binary.

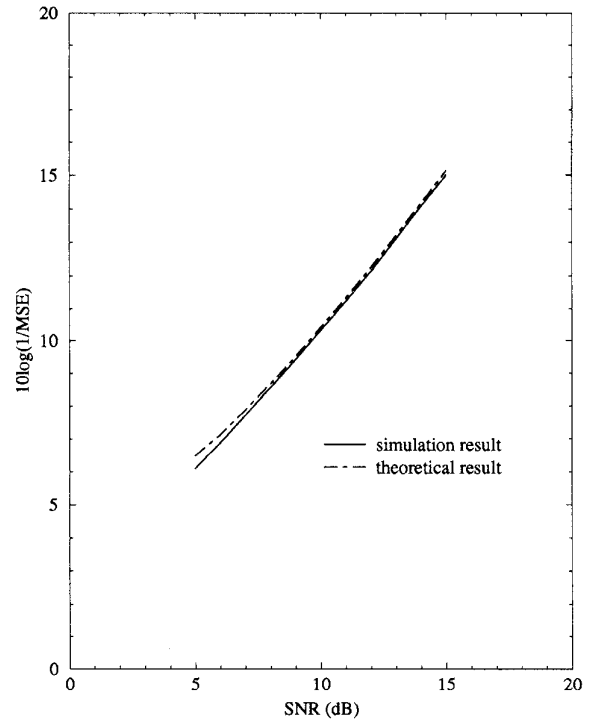


Fig. 6. MSE of the equalizer output when the channel input is of four-level.

To show this, we first find the cost near  $\{s_n\} = \{\frac{13}{31}(\delta[n-1] + \delta[n] + \delta[n+1])\}$ .

From (2.10), (2.12), and (2.13), we can find that

$$\Psi(y) = \begin{cases} \frac{1}{2}(|y| - 1)^2 & |y| < 1 \\ 0 & 1 \leq |y| \leq 3 \\ \frac{1}{2}(|y| - 3)^2 & |y| > 3. \end{cases} \quad (5.1)$$

Based on (2.5), for any  $s$  near  $s_o$ , we have

$$\begin{aligned} |y_n| < 1 & \quad \text{if } (a_{n+1}, a_n, a_{n-1}) \\ & \quad = \pm(3, -1, -1), \pm(-1, 3, -1), \\ & \quad \quad \pm(-1, -1, 3); \\ & \quad \quad \pm(1, 1, -1), \pm(1, -1, 1), \pm(-1, 1, 1); \\ |y_n| > 3 & \quad \text{if } (a_{n+1}, a_n, a_{n-1}) = \pm(3, 3, 3) \\ 1 < |y_n| < 3 & \quad \text{otherwise.} \end{aligned} \quad (5.2)$$

Direct calculation based on (5.2) can determine the cost function near  $s_o$  to be

$$\begin{aligned} f(s) = & \frac{65}{32} \sum_{n \neq -1, 0, 1} s_n^2 \\ & + \frac{1}{32} \sum_{n = -1, 0, 1} (31s_n^2 - 26s_n + 7). \end{aligned} \quad (5.3)$$

From (5.3), it is clear that  $s_o$  is local minimum of the SGA. Since the equalizer has no length restriction, it is a cost-dependent local minimum. Because of the symmetry of channel input and shift invariant nature of the cost function,  $\{\pm \frac{13}{31}(\delta[n-n_1] \pm \delta[n-n_2] \pm \delta[n-n_3])\}$  are all cost-dependent local minima of the SGA for all  $n_1 \neq n_2 \neq n_3$ .

Based on similar derivations, it can be shown that  $\{\pm \frac{37}{80}(\delta[n-n_1] \pm \delta[n-n_2] \pm \delta[n-n_3])\}$  are cost-dependent

local minima of DDE for all  $n_1 \neq n_2 \neq n_3$  since both the SGA and the decision feedback equalizers may have local minima even when the equalizers are double infinite. Hence, the center-tap initialization strategy [7], [12] cannot avoid the local convergence. This behavior will be demonstrated by computer simulations in the next section.

## VI. COMPUTER SIMULATION RESULTS

In this section, computer simulation examples will be presented to demonstrate the global convergence of the GA under Gaussian noise and the local convergence of the SGA.

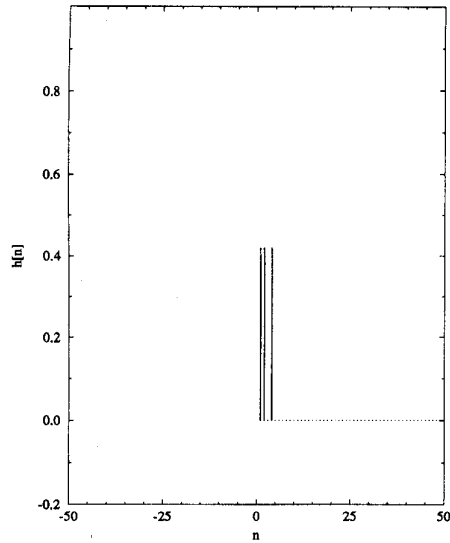
### A. Global Convergence of the Godard Equalizer Under Noise

The noisy channel considered in this example is given by

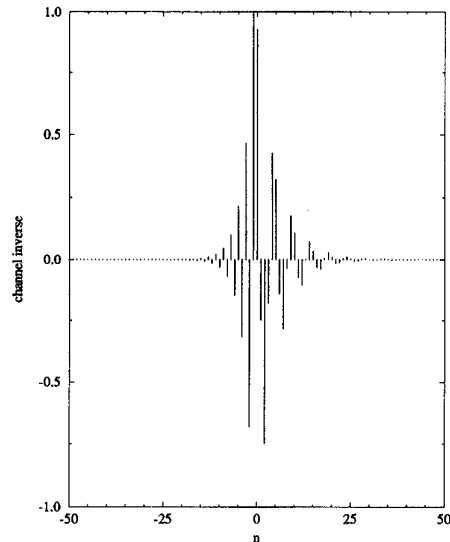
$$H(\omega) = \frac{e^{-j\omega} - 0.8}{1 - 0.8e^{-j\omega}} \quad (6.1)$$

which is an allpass channel with zero and pole at 1.25 and 0.8 respectively. The channel noise is white Gaussian. A 100-tap Godard equalizer with center-tap initialization is used in the system. To reduce the effect of gradient noise, step-size ( $\mu = 0.000002$ ) is set very small. The number of symbols used in the simulation is  $N = 100000$ .

When the channel input is binary, from (4.12), the MSE of the equalized system will be  $(1 + 1.25\eta)\eta$ . The simulated MSE and the comparative theoretical MSE are shown in Fig. 5. When the channel input is four-level, the MSE of the equalized system will be  $(1 - 0.9184\eta)\eta$ . Comparative results are shown in Fig. 6. It can be clearly seen that our analytical results are well supported by simulation examples.



(a)



(b)

Fig. 7. (a) Channel impulse response and (b) the channel inverse.

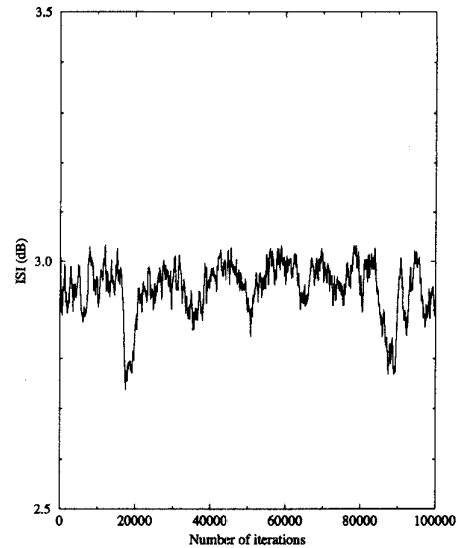
**B. Local Convergence of the Stop-and-Go Equalizer**

Consider a noiseless channel in PAM communication system is

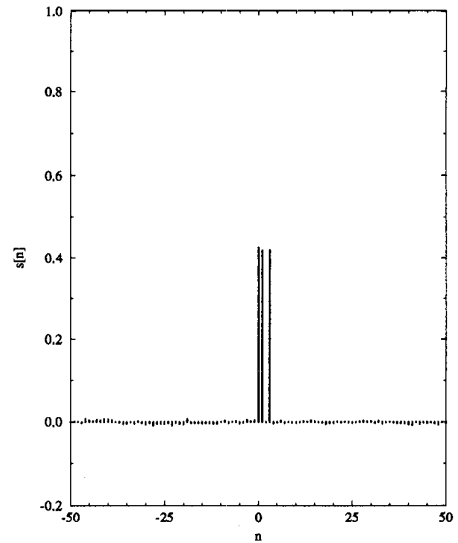
$$h(n) = \frac{13}{31}(\delta[n] + \delta[n - 1] + \delta[n - 3]). \quad (6.2)$$

The channel impulse response and the channel inverse are shown in Fig. 7. From Fig. 7(b), the channel inverse has about 20 significant coefficients. The channel input  $a_n$  is i.i.d., and uniform over the set  $\{-3, -1, 1, 3\}$ . A 100-tap stop-and-go equalizer with  $\beta = 2$  and small step size  $\mu = 0.0002$  is used. ISI is used to measure the performance of the equalized system, which is defined as

$$ISI = \frac{\sum_n s_n^2 - \max_n s_n^2}{\max_n s_n^2}. \quad (6.3)$$



(a)



(b)

Fig. 8. (a) ISI and (b) impulse response of the equalized system when the initial values of  $g_0$  is unit, and the rest of  $g_n$  are zero.

If the central-tap initial value, i.e.,  $\theta_0 = 1$  and  $\theta_n = 0$  for all  $n \neq 0$ , is used, the ISI and the impulse response of the equalized system are shown in Fig. 8.

From the simulation results, it can be seen that the equalizer is unable to remove the ISI and is, in fact, trapped in a cost-dependent local minimum. However, if the initial value is set such that  $g_0 = g_1 = 1$  and the rest of the coefficients are zero, the equalizer is able to remove the ISI, as shown by Fig. 9.

**VII. CONCLUSION**

In this paper, we present analytical results on the convergence of blind equalization algorithms. We have shown that every unconstrained blind equalization algorithm implemented with FIR filters possesses *length-dependent local minima*. Since cost-dependent local minima are common to all uncon-



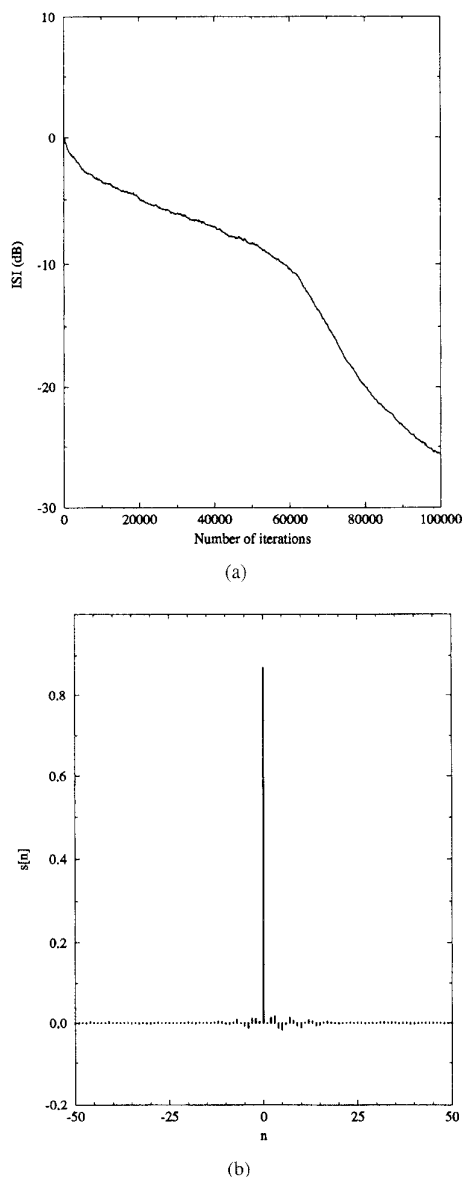


Fig. 9. (a) ISI and (b) impulse response of the equalized system when the initial values of  $g_0$  and  $g_1$  are unit, and the rest of  $g_n$  are zero.

strained blind adaptive equalizers, their regions of attraction should be used to determine the convergence performance of an algorithm. The local convergence property of blind equalization algorithms under the doubly infinite equalizer abstraction is cost dependent and is inherent to the cost function selection.

From the previous work in [3]–[5], [7], [12] and the discussion of this paper, it is known that the BGR algorithm [1], the Sato algorithm (SA) [18], the decision-directed equalizer (DDE) [8], and the stop-and-go algorithm (SGA) [17] all have cost-dependent local minima. On the other hand, the Godard algorithm (GA) [10], the Shalvi–Weinstein algorithm (SWA) [25], and standard cumulant algorithms (SCA's) [6] do not have cost-dependent local minimum. We also show that under high SNR, the Godard algorithm, the Shalvi–Weinstein

algorithm, and standard cumulant algorithms on global convergence do not amplify white channel noise since their ideal performance will be between the optimum Wiener equalizer and the channel inverse equalizer. Our analysis results are confirmed by computer simulations.

## 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. Automat. Contr.*, vol. AC-25, pp. 385–399, June 1980.
- [2] A. Benveniste and M. Goursat, "Blind equalizers," *IEEE Trans. Commun.*, vol. COM-32, pp. 871–882, Aug., 1982.
- [3] Z. Ding, R. A. Kennedy, B. D. O. Anderson, and C. R. Johnson, Jr., "Ill-convergence of Godard blind equalizers in data communication systems," *IEEE Trans. Commun.*, vol. 39, pp. 1313–1327, Sept. 1991.
- [4] ———, "On the (non)existence of undesirable equilibria of Godard blind equalizer," *IEEE Trans Signal Processing*, vol. 40, pp. 2425–2432, Oct. 1992.
- [5] ———, "Local convergence of the Sato blind equalizer and generalization under practical constraints," *IEEE Trans Inform. Theory*, vol. 39, pp. 129–144, Jan. 1993.
- [6] D. Donoho, "On minimum entropy deconvolution" in D. F. Findley, Ed., *Applied Time Series Analysis, II*. New York: Academic, 1981.
- [7] G. J. Foschini, "Equalization without altering or detecting data," *AT&T Tech. J.*, vol. 64, pp. 1885–1911, Oct., 1985.
- [8] D. A. George, R. R. Bowen, and J. R. Storey, "An adaptive decision feedback equalizer," *IEEE Trans. Commun. Tech.*, vol. COM-19, no. 3, June 1971.
- [9] G. B. Giannakis and J. M. Mendel, "Identification of nonminimum phase systems using via higher order statistics," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 37, pp. 360–377, 1989.
- [10] D. N. Godard, "Self-recovering equalization and carrier tracking in two-dimensional data communication systems," *IEEE Trans. Commun.*, vol. COM-28, pp. 1867–1875, Nov. 1980.
- [11] R. A. Kennedy, B. O. Anderson, and R. R. Bitmead, "Stochastic dynamics of blind decision feedback equalizer adaptation," in *Proc. Adaptive Syst. Contr. Signal Processing 1989 IFAC Symp.*, 1990, pp. 579–584.
- [12] Y. Li and Z. Ding, "Convergence analysis of finite length blind adaptive equalizer," *IEEE Trans. Signal Processing*, vol. 43, pp. 2120–2129, Sept. 1995.
- [13] ———, "Global convergence of fractionally spaced Godard equalizer," in *Proc. 26th Asilomar Conf. Signal, Syst. Comput.*, Oct. 1994.
- [14] O. Macchi and E. Eweda, "Convergence analysis of self-adaptive equalizers," *IEEE Trans. Inform. Theory*, vol. 30, pp. 161–176, Mar. 1984.
- [15] J. E. Mazo, "Analysis of decision-directed equalizer convergence," *Bell Syst. Tech. J.*, vol. 59, no. 10, pp. 1857–1876, Dec. 1980.
- [16] C. L. Nikias, "ARMA Bispectrum approach to nonminimum phase system identification," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 36, pp. 513–525, Apr. 1988.
- [17] G. Picchi and G. Prati, "Blind equalization and carrier recovery using a "stop-and-go" decision-directed algorithm," *IEEE Trans. Commun.* vol. COM-35, pp. 877–887, Sept. 1987.
- [18] Y. Sato, "A method of self-recovering equalization for multi-level amplitude modulation," *IEEE Trans. Commun.* vol. COM-23, pp. 679–682, June 1975.
- [19] C. K. Chan and J. J. Shynk, "Stationary points of the constant modulus algorithm for real Gaussian signals," *IEEE Trans. Signal Processing*, vol. 38, pp. 2176–2180, 1990.
- [20] J. R. Treichler and B. G. Agee, "A new approach to multipath correction of constant modulus signals," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-31, pp. 349–372, Apr. 1983.
- [21] J. R. Treichler, V. Wolff, and C. R. Johnson, Jr., "Observed misconvergence in the constant modulus adaptive algorithm," in *Proc. 25th Asilomar Conf. Signals, Syst. Comput.*, Pacific Grove, CA, 1991, pp. 663–667.
- [22] J. K. Tugnait, "Identification of linear stochastic systems via second and fourth-order cumulant matching," *IEEE Trans. Inform. Theory*, vol. IT-33, pp. 393–407, May 1987.
- [23] S. Verdú, B. D. O. Anderson, and R. A. Kennedy, "Blind equalization without gain identification," *IEEE Trans. Inform. Theory*, vol. 39, pp. 292–297, Jan. 1993.

- [24] R. A. Kennedy and Z. Ding, "Blind adaptive equalizers for QAM communication systems based on convex cost functions," *Opt. Eng.*, vol. 31, pp. 1189–1199, June 1992.
- [25] O. Shalvi and E. Weinstein, "New criteria for blind deconvolution of nonminimum phase systems (channels)," *IEEE Trans. Inform. Theory*, vol. 36, pp. 312–321, Mar. 1990.
- [26] J. K. Tugnait, O. Shalvi, and E. Weinstein, "Comments on 'New criteria for blind deconvolution of nonminimum phase systems (channels),' " *IEEE Trans. Inform. Theory*, vol. 38, pp. 210–213, Jan. 1992.
- [27] S. Vembu, S. Verdù, R. A. Kennedy, and W. Sethares, "Convex cost function in blind equalization," *IEEE Trans. Signal Processing*, vol. 42, pp. 1952–1960, Aug. 1994.



**Ye Li** was born in Jiangsu, China. He received the B. Eng and M. Eng degrees in 1983 and 1986, respectively, from the Department of Wireless Engineering, Nanjing Institute of Technology, Nanjing, China. He received the Ph.D. degree in 1994 from the Department of Electrical Engineering, Auburn University, Auburn, AL.

From March 1986 to May 1991, he was a teaching assistant and then a lecturer with the National Mobile Communication Laboratory, Southeast University, China. From September 1991 to September 1994, he was a research and teaching assistant with the Department of Electrical Engineering, Auburn University. From September 1994 to May 1996, he was a post-doctoral research associate with the Department of Electrical Engineering and Institute for Systems Research, University of Maryland, College Park. Since May 1996, he has been a Senior Technical Staff Member with Wireless Systems Research Department, AT&T Laboratories—Research, Holmdel, NJ. His general research interests include statistical signal processing, digital communications, and wireless mobile systems.



**K. J. Ray Liu** received the B.S. degree from the National Taiwan University in 1983 and the Ph.D. degree from the University of California, Los Angeles, in 1990, both in electrical engineering.

Since 1990, he has been with Electrical Engineering Department and Institute for Systems Research of University of Maryland, College Park, where he is an Associate Professor. His research interests span all aspects of signal processing with application to image/video, radar/sonar, wireless communications, and medical and biomedical technology. He has

published over 100 papers in these areas.

Dr. Liu has received numerous awards including the 1994 National Science Foundation Young Investigator Award, the IEEE Signal Processing Society's 1993 Senior Award (Best Paper Award), the George Corcoran Award in 1994 for outstanding contributions to electrical engineering education, and the Outstanding Systems Engineering Faculty Award in 1996 in recognition of outstanding contributions to interdisciplinary research, both from the University of Maryland. He is an Associate Editor of IEEE TRANSACTIONS ON SIGNAL PROCESSING, an editor of the *Journal of VLSI Signal Processing*, a member of the Design and Implementation of Signal Processing Systems Technical Committee, and a member of the Multimedia Signal Processing Technical Committee of IEEE Signal Processing Society.



**Zhi Ding** was born in Harbin, China. He received the B.Eng. degree in July 1982 from the Department of Radio Engineering, Nanjing Institute of Technology, Nanjing, China, and the M.A.Sc. degree from the Department of Electrical Engineering, University of Toronto, Toronto, Canada, in May 1987. He received the Ph.D. degree from the School of Electrical Engineering, Cornell University, Ithaca, NY, in August 1990.

He joined the faculty of Auburn University, Auburn, AL, in September 1990, where he is an associate professor in the Department of Electrical Engineering. He has held visiting positions with the Australian National University, the NASA Lewis Research Center, and the USAF Wright Laboratory. His main research interests include digital communications, signal processing, adaptive signal processing, blind equalization, and cyclostationary signal processing.