Ill-Convergence of Godard Blind Equalizers in Data Communication Systems

Zhi Ding, *Member, IEEE,* Rodney A. Kennedy, *Member, IEEE,*
Brian D. O. Anderson, *Fellow, IEEE,* and C. Richard Johnson, Jr., *Fellow, IEEE*

***Abstract*—Godard algorithms form an important class of adaptive blind channel equalization algorithms for QAM transmission. In this paper, the existence of stable undesirable equilibria for the Godard algorithms is demonstrated through a simple AR channel model. These undesirable equilibria correspond to local but nonglobal minima of the underlying mean cost function, and thus permit the ill-convergence of the Godard algorithms which are stochastic gradient descent in nature. Simulation results confirm predicted misbehavior. For channel input of constant modulus, it is shown that attaining the global minimum of the mean cost necessarily implies correct equalization. A criterion is also presented for allowing a decision at the equalizer as to whether a global or nonglobal minimum has been reached.**

I. Introduction

In band-limited data communications systems that are widely used today, each transmitted symbol is extended by the distortion of the analog channel over a much longer interval than its original duration, hence causing the undesirable intersymbol interference (ISI) effect. Adaptive equalizers are currently the primary devices used by the receiver to combat ISI introduced in telephone and radio channels.

Successful blind equalizers do not require a known training sequence for adequate initialization as conventional adaptive equalizers do. Blind equalization has very important applications in data transmission systems, particularly where sending a training sequence is unrealistic or costly. Among a number of blind equalizer schemes that have been introduced [1]–[3], a special class of blind equalizer proposed by Godard [3] is now well accepted and has been proposed for many applications, including the equalization of QAM (quadrature amplitude modulated) digital signals.

The Godard family of blind equalizers, indexed by a parameter p, generalizes the pioneering structure presented by Sato [2] (which is recovered as the special case when $p = 1$). The first indication that the blind Sato scheme can lead to false parameter convergence under adaptation and consequently to poor performance of the equalizer was given by Mazo [4]. There he showed in the somewhat special case of a noiseless channel having no ISI that an overparametrized Sato equalizer can lead to convergence to undesirable equilibria in the adaptation process. Macchi and Eweda [5] achieved similar results using a different analytical approach. A more significant contribution of [5] was to show that in the Sato scheme one has almost sure convergence to the ideal parameter setting once the eye diagram has opened. (See Kumar [6] for a similar result.) However, global convergence to one of the desirable equilibria (from an initially closed eye) of the Sato scheme has only been established for a nonpractical situation of an infinite number of equalizer parameters and for some specific continuous "symbol" distributions—the heuristic being that an alphabet of M-ary symbols may be approximated by a uniform distribution [1]. Nonetheless, this result (which is one of many found in [1]) is remarkable given the difficulty of the general problem of establishing global convergence only to one of the desired equilibria. In contrast, our work is directed towards showing that generally a practical Godard MA equalizer with finite parameters never has this ideal global convergence property when the symbol distribution is discrete, as in all QAM digital systems. Godard [3] also considered the problem of false equilibria, and showed that for an infinitely parameterized equalizer with infinite delay, false equilibria can exist but these were later shown by Foschini [7] to be locally unstable and are thus insignificant.

More recently Treichler *et al.* [8], [9] provided an alternative development and interpretation of a special case of the Godard family where $p = 2$, and labeled it the constant modulus algorithm or CMA. (While CMA has now been extended into a class of algorithms sometimes labeled CMA Version $p-q$ for various integers p and q, conventionally CMA refers to the original CMA Version 2–2, identical to the Godard $p = 2$ algorithm. In this paper we follow the same convention unless otherwise stated.) Stimulated by [8], [9], Johnson, Dasgupta, and Sethares [10] established local convergence properties of real CMA in a neighborhood of the desired equilibrium using averaging methods [11]. This work relates closely to open-eye convergence results found in [5] but uses a different analysis technique.

In this paper, after some background results (Section II) we show that in principle it is possible to test for the ill-convergence of any of the Godard schemes without explicit knowledge of the actual input sequence, which would appear essential to be able to do in practice. Then we establish

Paper approved by the Editor for Channel Equalization of the IEEE Communications Society. Manuscript received August 17, 1989; revised February 27, 1990. This work was supported by the National Science Foundation under Grant MIP-8608787 and MIP-8921003, by the Australian Telecommunications and Electronics Research Board, and by the ANU Centre for Information Science Research. This paper was presented at the 23rd Conference on Information Sciences and Systems, Baltimore, MD, March 1989.

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

R. A. Kennedy and B. D. O. Anderson are with the Department of Systems Engineering, Australian National University, Canberra, ACT 2601, Australia.

C. R. Johnson, Jr. is with the School of Electrical Engineering, Cornell University, Ithaca, NY 14853.

IEEE Log Number 9100828.

(Section III) the possibility of ill-convergence by deriving a set of undesired stable equilibria for the entire family of real Godard equalizers (when ISI is present), including the important CMA scheme. Our results are subsequently generalized (Section IV) to the family of complex Godard equalizers. Our techniques stress constructive procedures providing a clear picture of why these blind schemes fail, thus complementing the earlier results for the $p = 1$ Sato special case [4], [5]. Our results also stand in contrast to those of Godard [3] concerning false equilibria for the $p = 2$ case. His results presuppose an infinite length equalizer which can overparameterize the channel inverse; our (FIR) equalizers are constrained in length to exactly match that of the AR channel inverse such that we cannot encounter this particular type of false equilibrium. His theory is also incomplete in the sense that it only shows the existence of a particular class of false equilibria, which were later shown to be locally unstable for $p = 2$ by Foschini [7] (and thus of no practical concern), while the theory also fails to recognize the existence of the false equilibria exhibited in this paper. Finally, we mention an important work by Verdú [12], which motivated much of our work here.

II. Problem Formulation

A. Godard Equalizers

Fig. 1 shows the diagram of a data communication system where a Godard equalizer is used. A sequence of i.i.d., digital signals $\{a_k \in \mathbb{C}\}$ is sent by the transmitter through a channel exhibiting linear distortion thus generating the output sequence $\{x_k \in \mathbb{C}\}$. The objective of the equalizer is to recover by inversion (modulo a delay) the original sequence from the received sequence $\{x_k\}$. For a channel of $AR(n)$ structure, with parameter vector given by $\theta \triangleq [\theta_0\ \theta_1\ \cdots\ \theta_n]'$, i.e.,

$$\sum_{i=0}^{n} \theta_i x_{k-i} = a_k \tag{2.1}$$

one can use a $MA(m)$ equalizer, with parameter vector $\widehat{\theta}(k) \triangleq [\widehat{\theta}_0(k)\ \widehat{\theta}_1(k)\ \cdots\ \widehat{\theta}_m(k)]'$, to generate an output

$$z_k = \sum_{i=0}^{m} \widehat{\theta}_i(k) x_{k-i} \tag{2.2}$$

to help remove the channel ISI. In our analysis we first consider the case where there is negligible channel noise. Later in the analysis we introduce significant channel noise n_k which corrupts x_k.

The Godard class of algorithms [3] can be defined by specifying a positive cost function as follows:

$$J_p(z) \triangleq \frac{1}{2p}(|z|^p - R_p)^2, \qquad p \in \{1, 2, \cdots\}, \qquad z \in \mathbb{C} \tag{2.3}$$

where the dispersion constant R_p is defined in terms of the input signal $\{a_k\}$ by

$$R_p = \frac{E\left\{|a_k|^{2p}\right\}}{E\{|a_k|^p\}}. \tag{2.4}$$

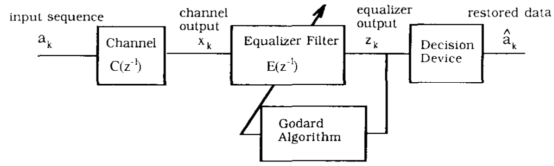


Fig. 1. Diagram of Godard channel equalization system.

Recall that $p = 1$ corresponds to Sato [2], and $p = 2$ to CMA (Version 2–2) [8]–[10].

The equalizer output (2.2) can be written $z_k = X_k'\widehat{\theta}(k)$ by taking $X_k \triangleq [x_k\ x_{k-1}\ \cdots\ x_{k-m}]'$ as the regressor vector. Then from (2.3) and (2.4), the adaptive algorithm takes the form

$$\widehat{\theta}(k+1) = \widehat{\theta}(k) - \mu \cdot \frac{\partial J_p(z_k)}{\partial \widehat{\theta}(k)} \tag{2.5a}$$

$$= \widehat{\theta}(k) - \mu \cdot z_k |z_k|^{p-2}(|z_k|^p - R_p)X_k^* \tag{2.5b}$$

where μ is the adaptation gain, $\frac{\partial}{\partial \widehat{\theta}(k)}$ denotes a complex vector differentiation operation (gradient) with respect to $\widehat{\theta}(k)$ (i.e., $\frac{\partial}{\partial \operatorname{Re}[\widehat{\theta}(k)]} + \frac{\partial}{\partial \operatorname{Im}[\widehat{\theta}(k)]}$), and X_k^* represents the complex conjugate of X_k. When $p = 1$ the update is defined as being zero at $z_k = 0$. In the case of single carrier QAM (i.e., PAM) where all the signals and parameters are real valued, the algorithm (2.5b) can be rewritten using $\operatorname{sgn}(z_k) = z_k|z_k|^{-1}$.

B. Global Minima and the Justification of the Mean Cost

In order to demonstrate that achieving the global minimum of the Godard mean cost necessarily implies the removal of ISI, we shall, in the remainder of Section II, restrict our attention to symbol alphabets with unit constant modulus $|a_k| = 1$ (i.e., PSK) where the dispersion constant R_p is unity.

The Godard algorithm (2.5) may be interpreted as a stochastic gradient descent procedure on the mean cost surface

$$\mathcal{J}_p\left(\widehat{\theta}(k)\right) \triangleq E\{J_p(z_k)\} = \frac{1}{2p}E\left\{(|z_k|^p - R_p)^2\right\} \tag{2.6}$$

where the expectation is to be taken over all $\{a_k\}$ data sequences. The mean cost $\mathcal{J}_p(\widehat{\theta}(k))$ clearly achieves the minimum value of zero when coefficients $\widehat{\theta}_i(k) = e^{j\phi}\theta_i\ \forall i$ for any ϕ, i.e., when the equalizer is correctly tuned [3]. What we shall argue now is that under a generally satisfied condition the converse holds, i.e., if the global minimum is achieved, then necessarily the equalizer is correctly tuned.

Result 1: Express the combined channel-equalizer transfer function (taking a_k to z_k) as

$$\frac{\sum_{i=0}^{m} \widehat{\theta}_i(k) z^{-i}}{\sum_{i=0}^{n} \theta_i z^{-i}} = h_0 + h_1 z^{-1} + h_2 z^{-2} + \cdots, \tag{2.7}$$

where $\{h_0, h_1, h_2, \cdots\}$ is BIBO stable with $h_i \in \mathbb{C}$, and suppose we have an arbitrary data sequence $\{\cdots, a_{k-2}, a_{k-1}, a_k, \cdots\}$ with symbols $a_{k-i} \in \mathbb{C} \cap \mathcal{A}$ taken from an alphabet $\mathcal{A} \triangleq \{\alpha, \beta, \gamma, \cdots\}$ of distinct QAM symbols, none of which is zero. Then if the equalizer output $z_k \triangleq \sum_{i=0}^{\infty} h_i a_{k-i}$ at time k satisfies a constant modulus property for all time and all data sequences, i.e.,

$$|z_k| = \left|\sum_{i=0}^{\infty} h_i a_{k-i}\right| = \rho > 0, \qquad \forall\{a_j\}, \quad \forall k \tag{2.8}$$

for some fixed constant $\rho \in \mathbb{R}$, then this implies perfect equalization, i.e.,

$$z_k = h_\delta a_{k-\delta}, \qquad \forall\{a_j\}, \quad \forall k \tag{2.9}$$

for some fixed decoding delay $\delta \in \mathbb{Z}_+$ independent of $\{a_j\}$ and k. In addition, the distinct symbols in alphabet $\mathcal{A}$ must have equal modulus, i.e., $|\alpha| = |\beta| = |\gamma| = \cdots$, and also the magnitude of the complex scalar h_δ must satisfy $|h_\delta| = \rho/|\alpha|$. □

For a proof of this and all subsequent results, see the Appendix. This result requires no statistical model of the elements of the $\{a_j\}$ sequence, cf. a similar style of result based on a comparison of the distributions of the symbols in $\mathcal{A}$ and the distribution of z_k, found in [1].

We now explain the relevance of this result to the justification of the choice of $J_p(z_k)$ when the dispersion constant R_p is unity. In this case $\rho = 1$ and we conclude that the input sequence $\{a_k\}$ and the equalizer output $\{z_k\}$ have the same constant modulus property $|a_k| = |z_k| = 1$, $\forall\{a_k\}$. Further, if we denote $h_\delta = e^{j\phi}$, then reflecting on (2.3) with $R_p = 1$, we may write

$$\frac{\sum_{i=0}^{m} \widehat{\theta}(k) z^{-i}}{\sum_{i=0}^{n} \theta_i z^{-i}} = e^{j\phi} z^{-\delta} \iff J_p(z_k) = 0 \quad \forall\, \{a_k\}. \tag{2.10}$$

Thus, the global minimization of the mean cost $\mathcal{J}_p(\widehat{\theta}(k))$ to zero (requiring $|z_k| = 1$ almost surely) is compatible with achieving a solution $\widehat{\theta}(k)$ that eliminates ISI.

Before moving on to look at local minima, we give a trivial relationship for later use between the user specified (scalar) cost function J_p (2.3) and the mean cost surface $\mathcal{J}_p$ (2.6) guiding adaptation.

Result 2: The scalar cost function $J_p(\cdot)$ can be recovered from the mean cost function $\mathcal{J}_p(\cdot)$ as follows:

$$\mathcal{J}_p(\gamma\theta) = \frac{1}{2p} E\left\{(|\gamma|^p - 1)^2\right\} = J_p(|\gamma|) = J_p(\gamma) \qquad \forall \gamma \in \mathbb{C} \tag{2.11}$$

where θ is the channel parameter vector satisfying $X_k'\theta = a_k$ with $|a_k| = 1$. □

Note: i) If we examine the mean cost $\mathcal{J}_p(\cdot)$ constrained to the one-dimensional (1-D) subspace $\{m\theta, m \in \mathbb{C}\}$ then we recover the original (user supplied) scalar cost function. ii) There will be other choices of cost $J(\cdot)$ for which $J(z_k) = 0$ if and only if $|z_k| = 1$. For example, there may be advantages in using a new $J(\cdot)$ with memory. However, Result 2 still holds for more general $J(\cdot)$.

C. Local Minima of the Mean Cost

We study the properties of the algorithm (2.5) by examining the incremental update when $R_p = 1$ (PSK). Following [3], [5], and [10], there will be an averaged equilibrium $\overline{\theta}$ when

$$E\left\{\frac{\partial J_p(z_k)}{\partial \widehat{\theta}(k)}\right\}\Bigg|_{z_k = X_k'\overline{\theta}} = E\Big\{z_k |z_k|^{p-2} \cdot (|z_k|^p - R_p) X_k^*\Big\}\Bigg|_{z_k = X_k'\overline{\theta}} = 0. \tag{2.12}$$

From (2.12), it is evident that $\overline{\theta} = 0$ and, for PSK (where $R_p = 1$)

$$\overline{\theta} \text{ such that } |z_k| = \left|X_k'\overline{\theta}\right| = 1, \qquad \forall k \tag{2.13}$$

are among the average equilibria for the Godard Algorithm given in (2.5). In fact, the equilibria satisfying (2.13) are the desired equilibria. Using Result 2 the undesirable $\overline{\theta} = 0$ will be seen to be unstable because the mean error surface $\mathcal{J}_p(\widehat{\theta}(k))$ has a local maximum at zero along the subspace $\{m\theta, m \in \mathbb{C}\}$ (since $J_p(m)$ has a local maximum at zero). If the equilibria of (2.13) are stable and if they represent *all* the existing stable equilibria, then the algorithm of (2.5) can achieve the objective of producing a constant modulus output z_k and eliminate the ISI.

Thus, the most important questions facing us are as follows. Are there any other equilibria besides the origin and the ones specified by (2.13)? Are they locally stable? We answer these two questions in the remaining sections of this paper. But first, we consider the crucial question of how measurements at the equalizer could be used to distinguish the convergence to undesirable equilibria from the convergence to desirable equilibria.

D. Testing for Undesirable Convergence

Note that when channel noise n_k is present or when there is imperfect modeling (e.g., if the equalizer is underparametrized), a precise condition like (2.8) or (2.9) cannot hold exactly. In this case the equilibria corresponding to desirable behavior may not be distinguishable by measurements of $J_p(\cdot)$ from any of the undesirable ones, because both will be nonzero. We therefore proceed to develop another test for distinguishing the correct equilibria from the undesirable ones (if they exist). While the test is developed for the zero noise, perfect modeling situation, it will evidently be robust in the face of modest departure from these ideal conditions.

Let y^* denote the complex conjugate, and y^H denote the complex conjugate transpose of vector $y \in \mathbb{C}^n$. The inner product of (2.12), given $R_p = 1$, with $\overline{\theta}$ yields the following necessary condition:

$$\overline{\theta}^H \cdot E\Big\{X_k'\overline{\theta}|X_k'\overline{\theta}|^{p-2}\Big(|X_k'\overline{\theta}|^p - 1\Big)X_k^*\Big\} = E\Big\{|z_k|^{2p}\Big\} - E\{|z_k|^p\} = 0. \tag{2.14}$$

This condition also holds for PAM systems (i.e., real signals).

To explain the condition of undesirable convergence, we will need the following two simple properties.

Property 1: Let the equalizer output z_k be a real (PAM) or complex random variable (QAM). If $E\{|z_k|^{2p}\} = E\{|z_k|^p\}$

for some $p \in \{1, 2, \cdots\}$, then either

$$\text{i)}\quad |z_k| = 1 \quad \text{almost surely, or} \tag{2.15a}$$

$$\text{ii)}\quad E\left\{|z_k|^{2p}\right\} < 1. \tag{2.15b}$$

□

Property 2: For the equalizer output z_k, if $E\left\{|z_k|^l\right\} < 1$, then $E\left\{|z_k|^m\right\} < 1, \forall\, 1 \le m \le l$. □

From (2.14), we know that the algorithm in (2.5) reaches equilibrium only if $E\left\{|z_k|^{2p}\right\} = E\left\{|z_k|^p\right\}$. Therefore, we can invoke Property 1 directly. Equation (2.15a) is exactly the ideal constant modulus objective of the blind algorithm, so that the associated equilibria are the *desired* ones. On the other hand, if the algorithm has reached an equilibrium but (2.15b) holds instead, then the resulting equilibrium is *undesirable* and by Property 2, we have

$$E\{|z_k|^m\} < 1. \qquad \forall\, 1 \le m \le 2p. \tag{2.16}$$

As a result, we know that if an equilibrium of the Godard algorithm ($\forall p \in \mathbb{Z}_+$) applied to an AR channel with unit modulus input is *undesirable*, then after convergence the following two statistical conditions are satisfied:

$$E\{|z_k|\} < 1 \qquad \text{the magnitude test} \tag{2.17a}$$

$$E\left\{|z_k|^2\right\} < 1 \qquad \text{power test.} \tag{2.17b}$$

By using time-averages, conditions (2.17a), (2.17b) help to indicate whether, after apparent algorithm convergence, it is an undesirable or a desirable equilibrium which has been reached. Testing in this manner does not require explicit knowledge of the input sequence $\{a_k\}$ other than that it has constant modulus.

Before proceeding to the following sections where we explicitly derive some undesired equilibria for the Godard algorithm, we first introduce the following property which will be useful later for characterizing the moments of x_k, the channel output (Fig. 1).

Property 3: If $E\{|x_k|\} \ge 1$ and m is an integer greater than 1, then either

i) $E\{|x_k|^m\} > E\left\{|x_k|^{m-1}\right\} > \cdots > E\left\{|x_k|^2\right\} > E\{|x_k|\}$, if $\Pr\{|x| \ne 1\} > 0$, or

ii) $E\{|x_k|^m\} = 1 \quad \forall m \in \mathbb{Z}_+$ if $|x_k| = 1$ almost surely.

□

III. Undesirable Equilibria for Real Godard Equalizers

In the preceding section, we have discussed the possible existence of some undesirable equilibria and presented conditions which are necessary for convergence to one of these equilibria. In this section, we shall display explicitly some undesirable, locally stable equilibria for a simple one-carrier (PAM) systems with a real channel and equalizer, and with zero-mean i.i.d. stationary input symbols, i.e., $0 < E\left\{|a_k|^2\right\} < \infty$ and $E\{a_k\} = 0$. These central results are valid for alphabets with arbitrary dispersion constants R_p (i.e., here and in Section IV, we do not assume $|a_k| = 1$). The case involving QAM transmission will be discussed separately in Section IV since for QAM we require $E\{a_k^2\} = 0$ [3], which is not possible in PAM transmission systems. Thus, the analysis involving PAM is not a special case of QAM system, as the stability results will show later.

A. Derivation of Undesirable Equilibria

Consider a particular real $AR(n)$ channel

$$\sum_{i=0}^{n} \theta_i z^{-i} = 1 + \alpha z^{-n}, \qquad \alpha \in \mathbb{R}, \qquad 0 < |\alpha| < 1 \tag{3.1}$$

and correspondingly, a general $MA(n)$ equalizer

$$\sum_{i=0}^{n} \widehat{\theta}_i(k) z^{-i} = \widehat{\theta}_0(k) + \widehat{\theta}_1(k) z^{-1} + \cdots + \widehat{\theta}_n(k) z^{-n}. \qquad \widehat{\theta}_i(k) \in \mathbb{R}. \tag{3.2}$$

Ideal equalization occurs when $\widehat{\theta}(k) = \pm\theta = \pm[1\ 0\ \cdots\ 0\ \alpha]'$. Let $\overline{\theta} = \left[\overline{\theta}_0\ \overline{\theta}_1\ \cdots\ \overline{\theta}_n\right]'$ denote an equilibrium of the real Godard algorithm (2.5) satisfying (2.12) that we seek explicitly.

Notice that the autoregressive channel output x_k from (3.1) can be written as

$$x_k = -\alpha x_{k-n} + a_k = \sum_{i=0}^{\infty} (-\alpha)^i a_{k-i\cdot n} \qquad |\alpha| < 1 \tag{3.3}$$

which means that for $k \in \mathbb{Z}_+$, the random variables x_k, $x_{k-1}, \cdots, x_{k-n+1}$ are independent of one another. Further since a_k has zero mean, we easily deduce $E\{x_k\} = 0$ for a stable channel ($|\alpha| < 1$).

As noted in [3], to solve for all the solutions (equilibria) of (2.12) based on (3.3) can be extremely difficult, even though the zero vector and $\pm\theta$ are known to be three of the existing solutions. Instead, we shall look specifically for equilibria of the form $\overline{\theta} = \left[0\ 0\ \cdots\ 0\ \overline{\theta}_n\right]'$, $\overline{\theta}_n \ne 0$, for which the equalizer output signal degenerates to $z_k = X'(k)\overline{\theta} = \overline{\theta}_n x_{k-n}$. Consequently, the equilibrium definition (2.12) becomes

$$E\left\{\left(|x_{k-n}|^p |\overline{\theta}_n|^p - R_p\right)|x_{k-n}|^{p-1}|\overline{\theta}_n|^{p-1} \cdot \operatorname{sgn}\left(\overline{\theta}_n x_{k-n}\right) x_{k-i}\right\} = 0, \qquad i = 0, 1, \cdots, n. \tag{3.4}$$

Due to the independence of x_k, $x_{k-1}, \cdots, x_{k-n+1}$, it follows that the above equations for $i = 1, 2, \cdots, n-1$ are automatically satisfied given $E\{x_{k-i}\} = 0$. Thus, only two equations (those for $i = 0$ and $i = n$) remain nontrivial in the $n+1$ equations of (3.4). Note that from (3.3), we can also get

$$E\left\{\operatorname{sgn}\left(\overline{\theta}_n x_{k-n}\right)|x_{k-n}|^m x_k\right\} = -\alpha E\left\{\operatorname{sgn}\left(\overline{\theta}_n x_{k-n}\right) \cdot |x_{k-n}|^m x_{k-n}\right\}. \tag{3.5}$$

Therefore, the two remaining equations in (3.4) for $i = 0$ and $i = n$ are related through a multiplicative factor of $-\alpha$, and thus we are down to a single equation

$$|\overline{\theta}_n|^p E\left\{\operatorname{sgn}(\overline{\theta}_n x_{k-n})|x_{k-n}|^{2p-1}x_{k-n}\right\} - R_p E\left\{\operatorname{sgn}(\overline{\theta}_n x_{k-n})|x_{k-n}|^{p-1}x_{k-n}\right\} = 0, \qquad \overline{\theta}_n \neq 0. \tag{3.6}$$

Multiplying through this equation by $\overline{\theta}_n$ reveals two nonzero solutions for this equation

$$\overline{\theta}_n = \pm\left(\frac{E\{|x_k|^p\}}{E\left\{|x_k|^{2p}\right\}} R_p\right)^{\frac{1}{p}}. \tag{3.7}$$

We therefore have arrived at a pair of undesirable equilibria for the real Godard algorithm

$$\overline{\theta} = \pm\left(\frac{E\{|x_k|^p\}}{E\left\{|x_k|^{2p}\right\}} R_p\right)^{1/p} [0 \quad 0 \quad \cdots \quad 0 \quad 1]'. \tag{3.8}$$

We can verify directly that the magnitude/power conditions of (2.17) hold at this pair of *undesirable* equilibria when $|a_k| = 1$ (PSK) for which $R_p = 1$. It follows from (3.8) that

$$E\left\{|z_k|^{2p}\right\} = E\left\{|X_k'\overline{\theta}|^{2p}\right\} = E\left\{|\overline{\theta}_n x_{k-n}|^{2p}\right\} = |\overline{\theta}_n|^{2p} E\left\{|x_{k-n}|^{2p}\right\} = \frac{E^2\{|x_k|^p\}}{E\left\{|x_k|^{2p}\right\}} \leq 1, \quad \text{(Cauchy-Schwartz Inequality)} \tag{3.9}$$

with equality if and only if $|x_k|^p = \rho$, $\rho \in \mathbb{R}$. But from Result 1 this equality cannot hold if $\alpha \neq 0$. Thus, we have $E\{|z_k|^{2p}\} < 1$, and from Property 2, the magnitude/power conditions of (2.17) are verified.

B. Stability Conditions of the Undesirable Equilibria

In discussing the stability (attractiveness) of the average equilibria of the adaptive algorithm, we shall use the well-known fact that the equilibrium $\overline{\theta}$ is locally stable (attractive) if the Hessian matrix

$$\boldsymbol{H}(\overline{\theta}) \triangleq \left.\frac{\partial^2 E\{J_p(z_k)\}}{\partial\widehat{\theta}^2(k)}\right|_{\widehat{\theta}(k)=\overline{\theta}} = E\left.\left\{\frac{\partial\nabla_{\widehat{\theta}(k)}J_p(k)}{\partial\widehat{\theta}(k)}\right\}\right|_{\widehat{\theta}(k)=\overline{\theta}} \tag{3.10}$$

is positive definite. The commutativity of differentiation and expectation in (3.10) follows by the smoothness of $J_p(\cdot)$ for $p \geq 2$. For $p = 1$ (Sato), difficulties arise: a Hessian can be defined but the commutativity fails and the analysis becomes formidable. We shall study the local stability of the *undesirable* equilibria only for $p \geq 2$.

At any given equilibrium $\overline{\theta}$ of (2.12), we have for $p \geq 2$

$$\boldsymbol{H}(\overline{\theta}) = E\left\{\frac{\partial}{\partial\widehat{\theta}(k)}\left[X_k(|z_k|^p - R_p)|z_k|^{p-1}\operatorname{sgn}(z_k)\right]\right\} = E\left\{\left[(2p-1)|z_k|^{2p-2} - R_p(p-1)|z_k|^{p-2}\right]X_kX_k'\right\}. \tag{3.11}$$

At either one of the pair of undesired equilibria given by (3.8), we get

$$\boldsymbol{H}(\overline{\theta}) = E\left\{\left[(2-p)|\overline{\theta}_n|^{2p-2}|x_{k-n}|^{2p-2} - R_p(p-1)|\overline{\theta}|^{p-2}|x_{k-n}|^{p-2}\right]X_kX_k'\right\}. \tag{3.12}$$

In analyzing the positive definiteness of this specific Hessian, we can reach the following stability condition.

Stability Condition 1: For the set of undesirable equilibria given by (3.8), the Hessian matrix of (3.11) is positive definite if and only if

$$(2p-1)\frac{E\left\{|x_k|^{2p-2}\right\}}{E\left\{|x_k|^{2p}\right\}} - (p-1)\frac{E\left\{|x_k|^{p-2}\right\}}{E\{|x_k|^p\}} > 0, \quad p \geq 2. \tag{3.13}$$

□

Therefore, the pair of undesirable equilibria for the real Godard equalizer are both *locally stable* if the condition (3.13) holds. This expression will prove essential in our example to follow.

C. Consequences of Ill-Convergence

The convergence of the pair of undesirable equilibria (3.8) makes the equalizer merely a scalar plus n-sample delay, thus no channel distortion (ISI) can be removed. Therefore, if the channel is sufficiently dispersive such that the initial eye diagram of x_k is closed (requiring an equalizer), then after convergence to these undesirable equilibria (3.8) a blind equalizer will not have opened the eye. In this case a nearest element decision device will not recover $\{a_k\}$ and poor error probability performance will ensue.

As a common practice, a decision-directed algorithm is used in conjunction with a blind equalizer—once the blind equalizer has reduced enough ISI, the decision-directed algorithm takes over. It is apparent, therefore, that if the Godard algorithm converges to any member of the set of undesirable equilibria which do not eliminate any ISI, subsequent switching to a decision-directed mode can prove futile when the channel is highly dispersive.

D. Examples of Stable Undesirable Equilibria

Consider the special Godard algorithm of $p = 2$, which is also known as the constant-modulus algorithm (CMA) [8]–[10]. It can be shown from the AR equation of (3.4) that

$$E\left\{|x_k|^2\right\} = \frac{E|a_k^2|}{1-\alpha^2} \quad \text{and} \quad E\left\{|x_k|^4\right\} = \frac{\left[E|a_k^4| - 3(E|a_k^2|)^2\right]}{1-\alpha^4} + 3\left(\frac{E|a_k^2|}{(1-\alpha^2)}\right)^2. \tag{3.14}$$

Consequently we have for CMA (where $p = 2$)

$$(2p-1)\frac{E\left\{|x_k|^{2p-2}\right\}}{E\left\{|x_k|^{2p}\right\}} - (p-1)\frac{E\left\{|x_k|^{p-2}\right\}}{E\{|x_k|^p\}}$$

$$= 3\,\frac{E\left\{|x_k|^2\right\}}{E\left\{|x_k|^4\right\}} - \frac{1}{E\left\{|x_k|^2\right\}}$$
$$= \frac{\left[3\left(E|a_k^2|\right)^2 - E|a_k^4|\right](1-\alpha^2)^2}{E|a_k^4|E|a_k^2|(1-\alpha^4) + 6\left(E|a_k^2|\right)^3\alpha^2} > 0,$$
$$\forall\ |\alpha| < 1 \quad \text{if } 3\left(E|a_k^2|\right)^2 > E|a_k^4|. \tag{3.15}$$

Hence, the stability condition (3.13) is satisfied as long as $3\left(E|a_k^2|\right)^2 - E|a_k^4| > 0$.

For the typical M-ary PAM signal $\{a_k\}$ uniformly distributed over M ($M > 1$, even) discrete levels $\{-(M-1)d, \cdots, -3d, -d, d, 3d, \cdots, (M-1)d\}$, the stability condition becomes

$$\begin{aligned}3\left(E|a_k^2|\right)^2 - E|a_k^4| &= 3\left[\frac{2}{M}d^2\sum_{i=1}^{M/2}(2i-1)^2\right]^2 \\ &\quad - \frac{2}{M}d^4\sum_{i=1}^{M/2}(2i-1)^4 \\ &= 3\left[\frac{1}{3}(M^2-1)d^2\right]^2 \\ &\quad - \frac{1}{15}(M^2-1)(3M^2-7)d^4 \\ &= \frac{2}{15}(M^4-1)d^4 > 0.\end{aligned} \tag{3.16}$$

Thus, it can be concluded that the undesirable equilibria

$$\overline{\theta} = \pm\left[0 \;\; 0 \;\; \cdots \;\; 0 \;\; \sqrt{\frac{1-\alpha^4}{3E(a_k^2)(1+\alpha^2) - 2R_2(1-\alpha^2)}}\right]', \tag{3.17}$$

resulted by substituting (3.14) into (3.8), are a pair of *stable* equilibria for CMA when the input is the uniformly distributed M-ary PAM signal.

As a concrete example, let the channel have the simplest $AR(1)$ form with $\alpha = 0.6$ whose input a_k and output x_k satisfies

$$x_k + 0.6x_{k-1} = a_k, \qquad \Pr(a_k = \pm 1) = 0.5 \tag{3.18}$$

and the equalizer parameter vector is simply $\widehat{\theta} = \left[\widehat{\theta}_0\ \widehat{\theta}_1\right]'$. The ideal equilibria for CMA are $\overline{\theta} = \pm[1\ 0.6]'$ and the pair of undesirable equilibria that were shown to be locally attractive are at $\overline{\theta} = \pm[0\ 0.5575]'$. Due to the stochastic gradient descent nature of CMA, the locally stable equilibrium of its adaptive algorithm corresponds to a local minimum of the mean cost $\mathcal{J}_p(\widehat{\theta}) = E\{J_p(X_k'\widehat{\theta})\}$. Thus, if the mean cost $\mathcal{J}_p$ is plotted as a function of $\widehat{\theta}_0$ and $\widehat{\theta}_1$, we should observe local minima at all these four equilibria and possibly at some further points.

In Fig. 2(a), the 3-D plot of $\mathcal{J}_2(\widehat{\theta})$ as a function of $\widehat{\theta}_1$ and $\widehat{\theta}_0$ is displayed. A contour plot of $\mathcal{J}_p$ is given in Fig. 2(b). They clearly show two global minima achieving $\mathcal{J}_2 = 0$ at the ideal equilibria $\overline{\theta} = \pm[1\ 0.6]'$. Furthermore, two additional local minima also appear at the undesirable equilibria $\overline{\theta} = \pm[0\ 0.5575]'$. Thus, our analytical results are verified in this

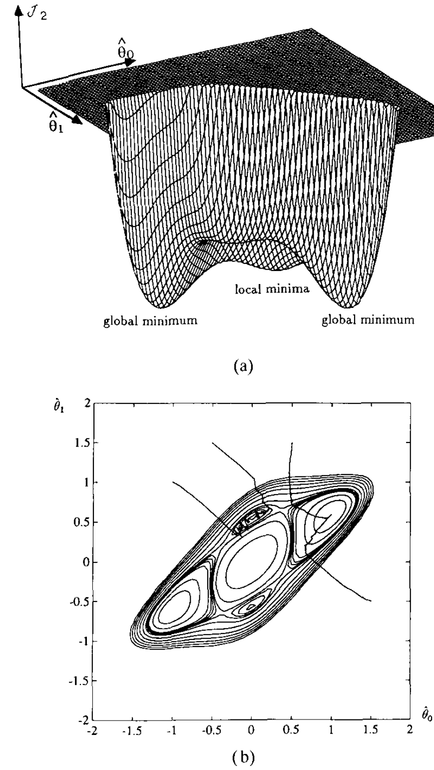
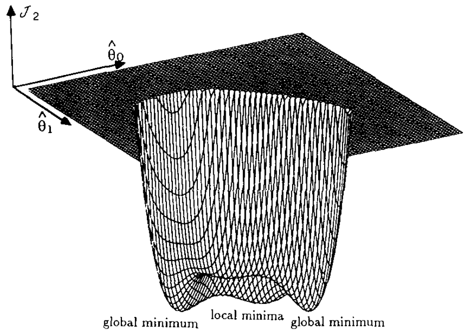


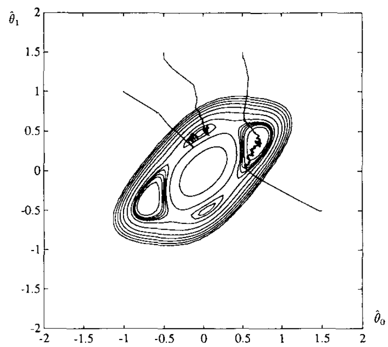
Fig. 2. Mean cost for binary transmission without channel noise under the $AR(1)$ channel. (a) Three-dimensional plot of the mean cost. (b) Contour plot of the mean cost and simulation trajectories.

simple example. Note, additionally, there are four saddle point solutions of (2.12) plus the maximum at the origin, leading to a total of nine "equilibria" (which theoretically exhausts all possibilities in this case). Notice that since the cost function is even, the contour of the mean cost is symmetric to the origin $[0\ 0]'$ as Fig. 2(b) clearly shows. In addition, the results of computer simulation (stepsize $\mu = 5 \times 10^{-3}$) of this equalization system are also presented through the plot of $\widehat{\theta}_0(k)$ versus $\widehat{\theta}_1(k)$ also shown in Fig. 2(b) under various initial conditions. The simulations show clearly that the pair of undesired equilibria $\overline{\theta}$ have considerable regions of attraction. The possibility for CMA to converge to either one of these undesired equilibria exists and is definitely not negligible.

We now maintain the same system except to let the input sequence $\{a_k\}$ be from an eight-level PAM, i.e., a_k uniformly distributed over $\{-7, -5, -3, -1, +1, +3, +5, +7\}$. We present the 3-D and the contour plot of the mean cost function $\mathcal{J}_2(\widehat{\theta})$ in Fig. 3(a) and (b), respectively. As we have shown analytically, while global minima exist at $\pm[1\ 0.6]'$, two additional local minima are present at $\pm[0\ 0.6830]'$. The 3-D and the contour plots bear the same features as in the case when the input is binary. Simulation results ($\mu = 10^{-5}$) of $\widehat{\theta}_0(k)$ and $\widehat{\theta}_1(k)$ initialized at the same four points as in Fig. 2(b) are also shown in Fig. 3(b) to have converged similarly to both the desired and the undesirable equilibria.



(a)



(b)

Fig. 3. Mean cost for 8-ary transmission without channel noise under the $AR(1)$ channel. (a) Three-dimensional plot of the mean cost. (b) Contour plot of the mean cost and simulation trajectories.

E. Channel Noise

Having established the possible ill-convergence of the Godard (CMA) algorithm under multilevel PAM input both analytically and experimentally, one important question remains unanswered, i.e., whether or not the presence of an additive noise n_k will change the surface of the cost function so that the undesirable equilibria either disappear or become very shallow such that they are almost unattractive to the algorithm. To answer this question, we will add in each of our two previous examples a white Gaussian additive noise n_k at x_k, independent of the input sequence $\{a_k\}$. Let the SNR at the channel output be 10 dB. We repeat the resulting 3-D and contour plots of the mean cost $\mathcal{J}_2(\widehat{\theta})$ in Figs. 4 and 5 for the binary transmission and the 8-ary transmission, respectively. (Of course, this requires a modification to the earlier analysis; details are omitted due to space limitations.) The contour plots are drawn at the same levels for both the noiseless and the noisy cases in order to illustrate the changes that occurred as a result of the additive noise n_k.

A comparison reveals that the shapes of the mean cost surface in both cases have been deformed by the inclusion of noise. Not only are the locations of the local and the global minima changed, but the depths of the two global minima have also been made shallower than for the noiseless channel. However, the regions of attraction for all the minima remain almost the same size. We can therefore expect simulations starting at a given location (initialization) to converge to the same type of (global and local) equilibria in both the noisy and the noiseless situations. Simulations under noise using the same stepsize ($\mu = 0.005$ and 10^{-5}, respectively) as in the noiseless cases are also illustrated in Figs. 4 and 5. They show the convergence of the four simulations to the same global and local minima as in the noiseless simulations. We, hence, can conclude that in general, the inclusion of small additive noises (SNR = 10 dB) at the channel output does not diminish the domain of attraction of the undesirable equilibria. While the existence of random noise makes it easier for the algorithm to escape from the local (undesirable) minima due to the possible large deviations, it also makes it easier for the algorithm to escape from the global (desirable) minima due to their reduced depth as a result of the noise effect. The hope that modest noise will decrease the chance of ill-convergence by the Godard algorithm is unsubstantiated.

F. Robustness

It is of interest to ask whether for an arbitrary AR channel of order n, rather than one of the special form (3.1), there can be *undesirable stable* equilibria. In fact, the same question can be posed for any channel model. This can be viewed as questioning the robustness of our results to perturbations in the model assumptions. A continuity argument indicates that for a channel close to (3.1), there will be undesirable equilibria close to (3.8). As one possible illustration for this, we simply draw the 3-D plot of $\mathcal{J}_p(\widehat{\theta})$ ($p = 2$) as a function of $\widehat{\theta}_0$ and $\widehat{\theta}_1$ for an MA channel consisting of the first three terms of the impulse response of our above example (3.18), i.e., the channel defined by

$$x_k = a_k - 0.6a_{k-1} + 0.36a_{k-2}. \tag{3.19}$$

$\{a_k\}$ is again i.i.d. uniform 8-level PAM and the SNR = 10 dB. A plot analogous to Fig. 2 is given in Fig. 6. It is apparent that undesirable equilibria remain. This result indicates the phenomena that we are studying are not dependent on the AR channel modeling assumption and it is only in the interest of clarity and simplicity that we isolated this analytically tractable situation.

Another robustness issue involves the assumption of the channel input being i.i.d. Although this assumption is common in most of the publications regarding (blind) adaptive equalization, ISI removal under correlated data input is an important problem. Our results of this paper can be generalized in principle to accommodate nonwhite inputs. In fact, in our computer simulations the input sequence is only pseudorandom and small correlation does exist between successive symbols.

Finally, we comment on the relationship between the undesirable equilibria developed here and in [3]. In [3], both for the case $p = 1$ and $p = 2$, a set of *undesirable* equilibria is defined in terms of the combined impulse response of the channel cascaded with the equalizer; in particular, the equalizer parameter vector that exhibits the *undesirable* equilibria of [3] is of infinite length, which is not a practical assumption. In addition, these undesirable equilibria have been shown to

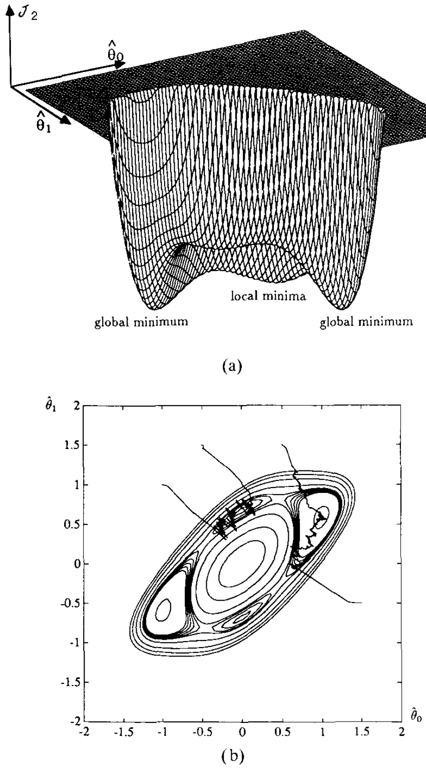


Fig. 4. Mean cost for binary transmission with channel SNR = 10 dB under the $AR(1)$ channel. (a) Three-dimensional plot of the mean cost. (b) Contour plot of the mean cost and simulation trajectories.

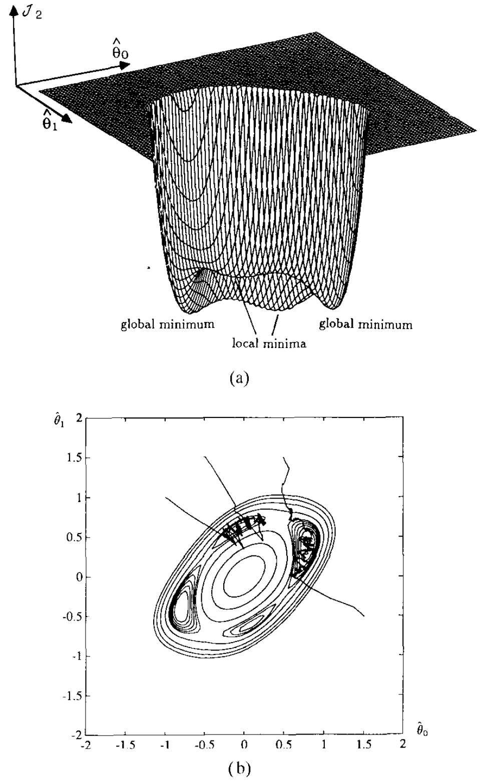


Fig. 5. Mean cost for 8-ary transmission with channel SNR = 10 dB under the $AR(1)$ channel. (a) Three-dimensional plot of the mean cost. (b) Contour plot of the mean cost and simulation trajectories.

be locally unstable for $p = 2$ and are hence not detrimental to the $p = 2$ Godard algorithm in practice. Observe that this excludes the *undesirable* equilibria that we have just developed here, which depend on the finite parametrization of the equalizer system. Hence, our results and those of Godard are complementary.

IV. Existence of Undesirable Equilibria for the Complex Godard Algorithm

In the previous section, we displayed a pair of stable undesirable equilibria with large domains of attraction for the real Godard algorithm. In this section, we shall analogously derive a set of undesirable equilibria for the complex (QAM) Godard algorithm, and establish their local stability condition which differs in form from the real case in a nontrivial way.

A. Derivation of Undesirable Equilibria

Consider now the complex case using the same $AR(n)$ channel as in (3.1) and (3.3), except $\alpha \in \mathbb{C}$, $|\alpha| < 1$. Suppose $\{a_k\}$ is an i.i.d., process which has zero mean, and the constellation has certain symmetries such that $E\{\mathrm{Re}^2(a_k)\} = E\{\mathrm{Im}^2(a_k)\} = E\{|a_k|^2\}/2$ and $E\{a_k^2\} = 0$ as in [3]. Notice again from (3.1) that the channel outputs $x_k, x_{k-1}, \cdots, x_{k-n+1}$ are independent identically distributed random variables satisfying $E\{x_k\} = E\{\mathrm{Re}(x_k)\} = E\{\mathrm{Im}(x_k)\} = 0$ since a_k has zero mean and $|\alpha| < 1$. For this $AR(n)$ channel, a general $MA(n)$ equalizer is required to eliminate ISI whose parameter vector is represented by (3.2) except the coefficients must be complex, i.e., $\widehat{\theta}_i(k) \in \mathbb{C}$. Let the output be given by $z_k = X_k'\widehat{\theta}(k)$.

Again we search equilibria of the form $\overline{\theta} = [0\,0 \cdots \overline{\theta}_n]'$ for which $z_k = X_k'\overline{\theta} = \overline{\theta}_n x_{k-n}$. In solving (3.1) the details mimic those of the real case and lead to the solution

$$\overline{\theta} = e^{j\psi}\left(\frac{E\{|x_k|^p\}}{E\{|x_k|^{2p}\}}R_p\right)^{1/p}[0 \quad 0 \quad \cdots \quad 1]', \quad \psi \in [0, 2\pi] \tag{4.1}$$

which defines a one-dimensional manifold of undesirable equilibria for the complex Godard equalizer. So again we see that a simple nontrivial channel leads to a set of equilibria for the complex blind Godard algorithm not associated with correct equalization. However, the question of stability for the complex algorithm requires a nontrivial extension from the real case and so it is here at this crucial point that we detail our analysis.

B. Stability Condition for the Undesired Equilibria

In order to examine the stability of this set of equilibria using the predominantly real stability theory, we need to

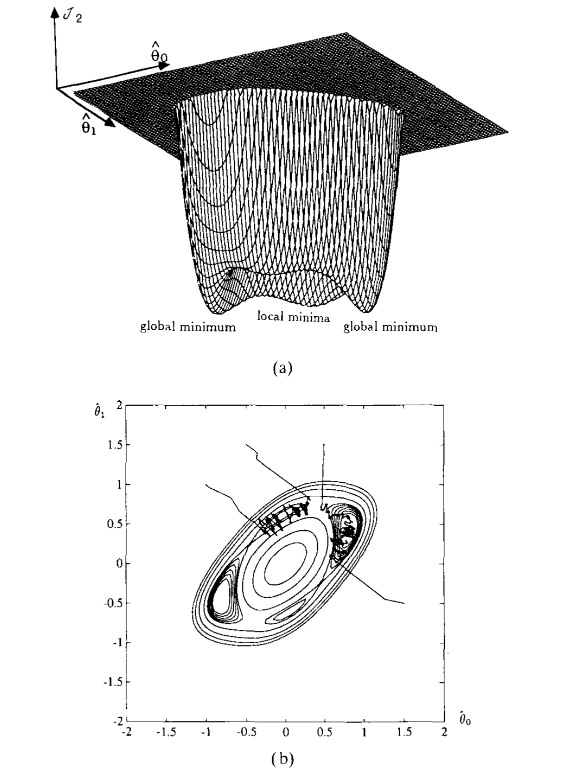


Fig. 6. Mean cost for 8-ary transmission with channel SNR = 10 dB under the $MA(2)$ channel. (a) Three-dimensional plot of the mean cost. (b) Contour plot of the mean cost and simulation trajectories.

consider a real arithmetic implementation of this complex algorithm (2.5). Write

$$R_k \triangleq \mathrm{Re}[x_k \quad x_{k-1} \quad \cdots \quad x_{k-n}]' \quad \text{and} \quad S_k \triangleq \mathrm{Im}[x_k \quad x_{k-1} \quad \cdots \quad x_{k-n}]'. \tag{4.2}$$

We now redefine the regressor matrix X_k and parameter vector $\widehat{\Theta}(k)$ as follows [13]:

$$X_k \triangleq \begin{bmatrix} R_k & S_k \\ -S_k & R_k \end{bmatrix}, \qquad \widehat{\Theta}(k) \triangleq \begin{bmatrix} \mathrm{Re}(\widehat{\theta}(k)) \\ \mathrm{Im}(\widehat{\theta}(k)) \end{bmatrix}. \tag{4.3}$$

Hence, the equalizer output can be represented as $Z_k \triangleq [\mathrm{Re}(z_k)\ \mathrm{Im}(z_k)]' = X_k'\widehat{\Theta}(k)$. With this new notation, we can rewrite the cost function $E\{J_p(z_k)\}$ (2.6) as

$$E\{J_p(Z_k)\} = \frac{1}{2p} E\left\{\left[\left(\widehat{\Theta}'(k)X_kX_k'\widehat{\Theta}(k)\right)^{p/2} - R_p\right]^2\right\}. \tag{4.4}$$

Further, the gradient (incremental algorithm update) is given by

$$\begin{aligned}\nabla_{\widehat{\Theta}(k)} J_p(Z_k) &= \left[\left(\widehat{\Theta}'(k)X_kX_k'\widehat{\Theta}(k)\right)^{p/2} - R_p\right] \\ &\quad \cdot \left(\widehat{\Theta}'(k)X_kX_k'\widehat{\Theta}(k)\right)^{p/2-1} X_kX_k'\widehat{\Theta}(k) \\ &= \left[(Z_k'Z_k)^{p/2} - R_p\right](Z_k'Z_k)^{p/2-1} X_kZ_k. \end{aligned} \tag{4.5}$$

The updating of the equalizer parameters is accomplished by

$$\widehat{\Theta}(k+1) = \widehat{\Theta}(k) - \mu\left[(Z_k'Z_k)^{p/2} - R_p\right](Z_k'Z_k)^{p/2-1} X_kZ_k. \tag{4.6}$$

This is actually a real arithmetic algorithm realization of the complex update scheme in (2.5). The average equilibrium $\overline{\Theta}$ of the Godard algorithm satisfies

$$E\left\{\left[\left(\widehat{\Theta}'(k)X_kX_k'\widehat{\Theta}(k)\right)^{p/2} - R_p\right] \left(\widehat{\Theta}'(k)X_kX_k'\widehat{\Theta}(k)\right)^{p/2-1} X_kX_k'\widehat{\Theta}(k)\right\}\Bigg|_{\widehat{\Theta}(k)=\overline{\Theta}} = 0. \tag{4.7}$$

Note that any particular solution $\widehat{\Theta}(k) = \overline{\Theta}$ of (4.7) has associated with it a manifold of equilibria

$$\overline{\Theta}_\psi \triangleq T_\psi \overline{\Theta}, \qquad \psi \in [0, 2\pi] \tag{4.8a}$$

where (let I_{n+1} denote the $n+1 \times n+1$ identity matrix)

$$T_\psi \triangleq \begin{bmatrix} I_{n+1}\cos\psi & I_{n+1}\sin\psi \\ -I_{n+1}\sin\psi & I_{n+1}\cos\psi \end{bmatrix} \tag{4.8b}$$

due to the cost function $J_p(Z_k)$ being insensitive to the phase.

For the complex $AR(n)$ channel (3.1), we may express the desired manifold of equilibria through

$$T_\psi[1 \quad 0 \quad \cdots \quad 0 \quad \alpha \quad 0 \quad 0 \quad \cdots \quad 0 \quad 0], \qquad \psi \in [0, 2\pi]. \tag{4.9}$$

Similarly, the manifold of undesirable equilibria (4.8) may be written

$$\overline{\Theta}_\psi \triangleq T_\psi \left(\frac{E\{|x_k|^p\}}{E\left\{|x_k|^{2p}\right\}} R_p\right)^{1/p} \cdot \Big[0 \quad 0 \quad \cdots \quad 0 \quad \underbrace{1}_{n+1\text{th}} \quad 0 \quad 0 \quad \cdots \quad 0 \quad 0\Big], \qquad \psi \in [0, 2\pi]. \tag{4.10}$$

In order to discuss the local attractiveness of this set of average equilibria, we need to consider the Hessian matrix $\boldsymbol{H}$ defined as

$$\begin{aligned}\boldsymbol{H}\left(\widehat{\Theta}\right) &\triangleq \frac{\partial^2 E\{J_p(Z_k)\}}{\partial \widehat{\Theta}^2(k)}\Bigg|_{\widehat{\Theta}(k)=\widehat{\Theta}} \\ &= E\left\{2\left[(p-1)\left(\widehat{\Theta}'X_kX_k'\widehat{\Theta}\right)^{p-2} - R_p(p/2-1) \cdot \left(\widehat{\Theta}'X_kX_k'\widehat{\Theta}\right)^{p/2-2}\right] X_kX_k'\widehat{\Theta}\widehat{\Theta}'X_kX_k'\right.\end{aligned}$$

$$+\left[\left(\widehat{\Theta}'X_kX_k'\right)^{p-1} - R_p\left(\widehat{\Theta}'X_kX_k'\widehat{\Theta}\right)^{p/2-1}\right] \cdot X_kX_k'\Bigg\}. \qquad (4.11)$$

Towards this end we present two key results.

Result 3: The one-dimensional smooth manifold of equilibria (4.8) is locally attractive if at every point on the manifold:

a) the $2(n+1)\times 2(n+1)$ Hessian matrix (4.11) is nonnegative definite with rank $2n+1$; and

b) the eigenvector corresponding to the zero eigenvalue is tangent to the manifold. □

Result 4: For the one-dimensional manifold of parameter vectors specified by

$$\overline{\Theta}_\psi = T_\psi \overline{\Theta}_0, \qquad \psi \in [0, 2\pi] \qquad (4.12)$$

$\boldsymbol{H}(\overline{\Theta}_\psi)$ has the same eigenvalues for all $\psi \in [0,2\pi]$. At every value of ψ, the eigenvector $v_0(\overline{\Theta}_\psi)$ corresponding to the zero eigenvalue of $\boldsymbol{H}(\overline{\Theta}_\psi)$ is always tangent to the manifold. □

In our study of the Godard equalizer used for complex $AR(n)$ channel equalization, the actual $2(n+1)\times 2(n+1)$ Hessian evaluated at $\overline{\Theta}_0$ is given by

$$\begin{aligned}\boldsymbol{H}(\overline{\Theta}_0) = 2E\Bigg\{&\Bigg[(p-1)\left|\overline{\Theta}_o\right|^{2p-4}|x_{k-n}|^{2p-4} - R_p(p/2-1) \\ &\cdot \left|\overline{\Theta}_0\right|^{p-4}|x_{k-n}|^{p-4}\Bigg] X_kX_k'\overline{\Theta}_0\overline{\Theta}_0'X_kX_k'\Bigg\} \\ &+ E\Bigg\{\Bigg[\left|\overline{\Theta}_0\right|^{2p-2}|x_{k-n}|^{2p-2} - R_p\left|\overline{\Theta}_0\right|^{p-2} \\ &\cdot |x_{k-n}|^{p-2}\Bigg] X_kX_k'\Bigg\}. \end{aligned} \qquad (4.13)$$

Stability Condition 2: For the complex $AR(n)$ channel defined by (3.1), the above $2(n+1)\times 2(n+1)$ Hessian matrix $\boldsymbol{H}(\overline{\Theta}_0)$ is nonnegative definite with rank $2n+1$ if and only if

$$\frac{E\left\{|x_k|^{2p-2}\right\}}{E\left\{|x_k|^{2p}\right\}} > \frac{1}{2}\,\frac{E\left\{|x_k|^{p-2}\right\}}{E\{|x_k|^p\}}. \qquad (4.14)$$

In addition, its eigenvector v_0 corresponding to zero eigenvalue is the tangent of the manifold $\overline{\Theta}_\psi$ at $\psi = 0$. □

Thus, if (4.14) holds, the Hessian matrix $\boldsymbol{H}(\widehat{\Theta})$ has rank $2n+1$ on the manifold of undesired equilibria $\overline{\Theta}_\psi \triangleq T_\psi\overline{\Theta}_0$, and at each point $\overline{\Theta}_\psi$, its eigenvector corresponding to zero eigenvalue is tangent to the manifold $\overline{\Theta}_\psi$. Therefore, the conditions of Result 3 are satisfied and the set of undesirable equilibria $\overline{\Theta}_\psi$ is locally stable if (4.14) holds. This is a sufficient condition for the set of undesired equilibria given by (4.10) to be locally stable.

As an example of applying (4.14) we test the stability of the undesirable equilibria for the $(p=2)$ CMA algorithm. From the autoregressive equation of (3.3) and based on the assumptions about the complex input sequence $\{a_k\}$ [3], we have

$$E\left\{|x_k|^2\right\} = \frac{E|a_k|^2}{1-|\alpha|^2}; \quad \text{and}$$

$$E\left\{|x_k|^4\right\} = \frac{E|a_k^4| - 2(E|a_k^2|)^2}{1-|\alpha|^4} + 2\left(\frac{E|a_k^2|}{1-|\alpha|^2}\right)^2. \qquad (4.15)$$

As a result, it follows that for $p = 2$

$$\begin{aligned}&\frac{E\left\{|x_k|^{2p-2}\right\}}{E\left\{|x_k|^{2p}\right\}} - \frac{1}{2}\,\frac{E\left\{|x_k|^{p-2}\right\}}{E\{|x_k|^p\}} = \frac{E\left\{|x_k|^2\right\}}{E\left\{|x_k|^4\right\}} - \frac{1}{2E\left\{|x_k|^2\right\}} \\ &= \frac{\left[2(E|a_k^2|)^2 - E|a_k^4|\right]\left(1-|\alpha|^2\right)^2\left(1+|\alpha|^2\right)}{2E|a_k^2|\left[E|a_k^4|\left(1-|\alpha|^2\right) + 4(E|a_k^2|)^2|\alpha|^2\right]} > 0, \\ &\qquad \forall |\alpha| < 1 \quad \text{if } 2(E|a_k^2|)^2 - E|a_k^4| > 0. \end{aligned} \qquad (4.16)$$

Consider the typical QAM where a_k is uniformly distributed over the constellation in which $\mathrm{Re}\{a_k\}$ and $\mathrm{Im}\{a_k\}$ both can take values from the same discrete levels $\{-(M-1)d, \cdots, -3d, -d, d, 3d, \cdots, (M-1)d\}$ (e.g., CCITT rectangular 16-QAM and 64-QAM). Then, it easily follows that

$$\begin{aligned}2(E|a_k^2|)^2 - E|a_k^4| &= 8\left(E\left|\mathrm{Re}\{a_k\}^2\right|\right)^2 - 2E\left|\mathrm{Re}\{a_k\}^4\right| \\ &\quad - 2E\left(\left|\mathrm{Re}\{a_k\}^2\right|\left|\mathrm{Im}\{a_k\}^2\right|\right) \\ &= 2\left[3\left(E\left|\mathrm{Re}\{a_k\}^2\right|\right)^2 - E\left|\mathrm{Re}\{a_k\}^4\right|\right] \\ &= \frac{4}{15}(M^4-1)d^4 > 0. \end{aligned} \qquad (4.17)$$

Therefore, for the complex CMA algorithm, the set of undesired equilibria given by (4.10) are locally stable when the typical QAM constellation as described above are used for which Stability Condition 2 holds. In fact, the stability condition also holds for inputs from the CCITT V.29 16-QAM constellation.

The remarks of Sections III-E and -F regarding robustness of the results also apply to the results of this section.

V. Conclusion

Aspects of the convergence behavior of Godard blind equalizers for QAM signal reception have been studied. It is shown that the global minimum of the mean cost under constant modulus input corresponds to the elimination of ISI by the equalizer setting. The Godard algorithms can be interpreted as an attempt to minimize a certain mean cost through adaptation of equalizer parameters, and through this view the existence of undesirable, i.e., incorrect, equilibria for Godard algorithms was demonstrated. It is shown with a simple system setup that these particular undesirable equilibria are locally stable for many useful and popular QAM constellations. Our examples and simulation confirm our findings. While in the original setting the channel noise was assumed zero and the channel was assumed to by an AR type, our examples and simulations show that violation of either condition does not result in the disappearance of the undesirable equilibria, contradicting many wishful conjectures. We hope that our findings will stimulate more research on many of the important and interesting issues of blind equalization, including searching for potential algorithms of global asymptotic optimality.

Our results proved the existence of ill-convergence by the Godard algorithm. They lack the generality one might desire due to the assumption of AR channel models. It will be of great interest and importance to extend these results into other, more general types of channel models in support of our simulation results. Although (3.8) also represents *undesirable* equilibria for the algorithm of $p = 1$ (Sato), a closed-form expression for their location and local stability condition have not been successfully determined due to the difficulty caused by the discontinuity of the gradient of the cost function at the origin. However, simulations similar to that in Fig. 6 reveal the local stability of the undesirable equilibria of (3.8) for the simple example. Further work is needed to derive the associated stability condition. Also, our convergence test derived in Section II is derived for channel inputs of constant modulus. Further study is needed to establish a similar test for a more general class of QAM inputs. The main results of this paper motivate the need for further work in the development of equalizer initialization tactics that can help to prevent Godard algorithms from exhibiting ill-convergence to undesirable equilibria.

APPENDIX
PROOFS OF RESULTS, PROPERTIES, AND STABILITY CONDITIONS

Proof of Result 1: Suppose $h_i \neq 0$ and $h_j \neq 0$ for at least two integers $i \neq j$. Consider four different data sequences labeled $\{a_k^{(1)}\}$, $\{a_k^{(2)}\}$, $\{a_k^{(3)}\}$, $\{a_k^{(4)}\}$, where $a_{k-m}^{(1)} = a_{k-m}^{(2)} = a_{k-m}^{(3)} = a_{k-m}^{(4)} = a_k$, $\forall\, m \neq i,\, m \neq j$ and

$$\begin{cases} a_{k-i}^{(1)} = \alpha \\ a_{k-j}^{(1)} = \alpha \end{cases};\quad \begin{cases} a_{k-i}^{(2)} = \alpha \\ a_{k-j}^{(2)} = \beta \end{cases};\quad \begin{cases} a_{k-i}^{(3)} = \beta \\ a_{k-j}^{(3)} = \alpha \end{cases};\quad \begin{cases} a_{k-i}^{(4)} = \beta \\ a_{k-j}^{(4)} = \beta \end{cases}. \tag{A.1.1}$$

With $\alpha \neq \beta$, it then follows that $z_k^{(1)} = R + h_i\alpha + h_j\alpha$ where $R = \sum_{m\neq i,j} h_m a_{k-m}$. Similarly, we have $z_k^{(2)} = R + h_i\alpha + h_j\beta$, $z_k^{(3)} = R + h_i\beta + h_j\alpha$, and $z_k^{(4)} = R + h_i\beta + h_j\beta$. Here $z_k^{(1)}$, $z_k^{(2)}$, $z_k^{(3)}$, and $z_k^{(4)}$ can be seen as four different vectors in the complex plane. It is clear that

$$z_k^{(1)} - z_k^{(2)} = z_k^{(3)} - z_k^{(4)} = (\alpha - \beta)h_j \neq 0 \tag{A.1.2a}$$

$$z_k^{(1)} - z_k^{(3)} = z_k^{(2)} - z_k^{(4)} = (\alpha - \beta)h_i \neq 0. \tag{A.1.2b}$$

Thus, $z_k^{(1)}$, $z_k^{(2)}$, $z_k^{(3)}$, and $z_k^{(4)}$ must be four vertices of a parallelogram in the complex plane (or they can also lie in a line which is also a special parallelogram). But since we have by hypothesis that $|z_k^{(1)}| = |z_k^{(2)}| = |z_k^{(3)}| = |z_k^{(4)}| = \rho$, it is evident that this parallelogram has to be a rectangle centered at the origin. Hence we can conclude that $R = 0$.

Since $R = 0$, it follows from $|z_k^{(1)}| = |z_k^{(4)}|$ that $|\alpha| = |\beta|$, i.e., $\alpha = \beta e^{j\phi}$ where $\phi \neq 2n\pi$, $n \in \mathbb{Z}_+$. Consequently, from (A.1.2) we have

$$|\alpha||h_i + h_j| = |\alpha||h_i + e^{j\phi}h_j| = |\alpha||h_i + e^{-j\phi}h_j|, \quad \phi \neq 2n\pi \tag{A.1.3}$$

which means either $h_i = 0$ or $h_j = 0$. The last equation contradicts our assumption of $h_i \neq 0$ and $h_j \neq 0$. Thus, we can conclude that there exists only one $h_\delta \neq 0$ for some $\delta \in \mathbb{Z}_+$. Now from

$$|z_k| = \left|\sum_{j=0}^{\infty} h_j a_{k-\delta}\right| = |h_\delta a_{k-\delta}| = \rho, \qquad \forall\, a_{k-\delta} \in \mathcal{A} \tag{A.1.4}$$

it is clear that all symbols $\mathcal{A} \triangleq \{\alpha, \beta, \gamma, \cdots\}$ satisfy $|\alpha| = |\beta| = |\gamma| = \cdots$, and thus $|h_\delta| = \rho/|\alpha|$. □

Proof of Property 1: By the Cauchy–Schwartz inequality, it is known that for random variables x, y, $E^2\{xy\} \leq E\{x^2\}E\{y^2\}$ with strict equality if and only if $x = \lambda y$, $\lambda \in \mathbb{R}$. Now let $x = |z|^p$ and $y = 1$. Then $E^2\{|z|^p\} \leq E\{|z|^{2p}\}$ in which the equality holds if and only if $|z|^p = \lambda$, for some $\lambda \in \mathbb{R}$. But by the hypothesis $E^2\{|z|^p\} \leq E\{|z|^{2p}\} = E\{|z|^p\}$ which gives $E\{|z|^p\} \leq 1$ with strict equality if and only if $|z|^p = \lambda = 1$ or $|z| = 1$ (except on a set of measure zero). □

Proof of Property 2: (Outline) It is easy to verify that the real polynomial $mx^l - lx^m - m + l$ for $x \geq 0$ and $l \geq m \geq 1$ has a global minimum at $x = 1$, and therefore takes only positive values. Thus, (letting $x = |z|$), $m(|z|^l - 1) \geq l(|z|^m - 1)$, $\forall z \in \mathbb{C}$, $\forall l \geq m \geq 1$. Therefore, by hypothesis and taking expectations $0 > m(E\{|z|^l\} - 1) \geq l(E\{|z|^m\} - 1)$ which readily leads to the desired result. □

Proof of Property 3: (Outline) It is easy to verify for arbitrary i, $(|x|^i - 1)(|x| - 1) \geq 0$, $\forall\, x \in \mathbb{C}$ with equality only when $|x| = 1$. Therefore, taking expectations we obtain $E\{|x|^{i+1}\} - E\{|x|^i\} - E\{|x|\} + 1 \geq 0$ with equality only if $|x| = 1$ almost surely. The desired result now easily follows from the hypothesis. □

Proof of Stability Condition 1: Write the Hessian (3.12) as

$$\boldsymbol{H}(\overline{\theta}) = E\{f(x_{k-n})X_kX_k'\} = E\left\{ f(x_{k-n}) \begin{bmatrix} x_k^2 & x_kx_{k-1} & \cdots & x_kx_{k-n} \\ x_kx_{k-1} & x_{k-1}^2 & \cdots & x_{k-1}x_{k-n} \\ \vdots & \vdots & \ddots & \vdots \\ x_kx_{k-n} & x_{k-1}x_{k-n} & \cdots & x_{k-n}^2 \end{bmatrix} \right\} \tag{A.5.1}$$

where

$$f(x_{k-n}) \triangleq (2p-1)|\overline{\theta}_n|^{2p-2}|x_{k-n}|^{2p-2} - (p-1)|\overline{\theta}_n|^{p-2}|x_{k-n}|^{p-2}. \tag{A.5.2}$$

Using the independence between x_{k-i}, $i = 1, 2, \cdots, n-1$ and

x_{k-n} [and hence $f(x_{k-n})$], we get (A.5.3) as shown below. From the independence of a_k and x_{k-n} and the channel characteristics, it is positive definite if and only if

$$E\{x_{k-n}^2 f(x_{k-n})\} = |\overline{\theta}_n|^{p-2} E\left\{(2p-1)|\overline{\theta}_n|^p |x_{k-n}|^{2p} - R_p(p-1)|x_{k-n}|^p\right\} = pR_p|\overline{\theta}_n|^{p-2}E\{|x_k|^p\} > 0 \quad \text{(A.5.4)}$$

which is always satisfied and

$$E\{f(x_{k-n})\} = |\overline{\theta}_n|^{p-2} E\left\{(2p-1)|\overline{\theta}_n|^p |x_{k-n}|^{2p-2} - R_p(p-1)|x_{k-n}|^{p-2}\right\} = R_p|\overline{\theta}_n|^{p-2}E\left\{(2p-1)\frac{E\{|x_k|^p\}}{E\{|x_k|^{2p}\}}|x_k|^{2p-2} - (p-1)|x_k|^{p-2}\right\} > 0. \quad \text{(A.5.5)}$$

Therefore, $\boldsymbol{H}(\overline{\theta})$ given in (A.5.1) will be positive definite if and only if the scalar expectation in (A.5.5) is positive. This condition is equivalent to (3.13). □

Proof of Result 3: From (2.5), the average system is

$$\widehat{\Theta}(k+1) = \widehat{\Theta}(k) - \mu \cdot E\left\{\frac{\partial J_p(z_k)}{\partial \widehat{\Theta}(k)}\right\}. \quad \text{(A.6.1)}$$

By denoting $g(\widehat{\Theta}(k)) = E\{\nabla_{\widehat{\Theta}(k)} J_p(k)\}$, the equilibrium of the average system is given by

$$g\left(\widehat{\Theta}(k)\right) = 0. \quad \text{(A.6.2)}$$

Let $\overline{\Theta}$ be a solution to (A.6.2) which is hence an average equilibrium. Then the average system can be linearized in the neighborhood of $\overline{\Theta}$ (Euclidean norms are understood)

$$\widehat{\Theta}(k+1) - \overline{\Theta} = \widehat{\Theta}(k) - \overline{\Theta} + \mu g(\overline{\Theta}) + \mu \frac{\partial g\left(\widehat{\Theta}(k)\right)}{\partial \widehat{\Theta}(k)}\Bigg|_{\overline{\Theta}} \cdot \left[\widehat{\Theta}(k) - \overline{\Theta}\right] + O\left(\left\|\widehat{\Theta}(k) - \overline{\Theta}\right\|^2\right)$$
$$= \left[I - \mu \boldsymbol{H}(\overline{\Theta})\right] \cdot \left[\widehat{\Theta}(k) - \overline{\Theta}\right] + O\left(\left\|\widehat{\Theta}(k) - \overline{\Theta}\right\|^2\right). \quad \text{(A.6.3)}$$

Since $\overline{\Theta}_\psi$ of (4.8) defines a 1-manifold of equilibria for the average system, we shall define a special *Lyapunov Function*

$$V(k) \triangleq \inf_\psi \left\|\widehat{\Theta}(k) - \overline{\Theta}_\psi\right\|. \quad \text{(A.6.4)}$$

For $\widehat{\Theta}(k)$ in the neighborhood of the one-dimensional manifold $\overline{\Theta}_\psi$, we can find a $\overline{\Theta}_{\psi'}$ on the 1-manifold that is closest to $\widehat{\Theta}(k)$. Clearly,

$$V(k) = \left\|\widehat{\Theta}(k) - \overline{\Theta}_{\psi'}\right\| \quad \text{(A.6.5)}$$

and $\widehat{\Theta}(k) - \overline{\Theta}_{\psi'}$ is orthogonal to the tangent plane of the manifold $\overline{\Theta}_\psi$ at ψ', otherwise $\overline{\Theta}_{\psi'}$ cannot achieve the infinum of (A.6.4). The eigenvectors of $\boldsymbol{H}(\overline{\Theta}_{\psi'})$ are all orthogonal, since $\boldsymbol{H}(\overline{\Theta}_{\psi'})$ is symmetric, and so (under the hypotheses) $\widehat{\Theta}(k) - \overline{\Theta}_{\psi'}$ must be in the space spanned by those eigenvectors of $\boldsymbol{H}(\overline{\Theta}_{\psi'})$ corresponding to the positive eigenvalues. Thus, for μ small, $\left[I - \mu \boldsymbol{H}(\overline{\Theta}_{\psi'})\right]$ is a contraction for $\left[\widehat{\Theta}(k) - \overline{\Theta}_{\psi'}\right]$ and we have

$$\left\|\left[I - \mu \boldsymbol{H}(\overline{\Theta}_{\psi'})\right]\left[\widehat{\Theta}(k) - \overline{\Theta}_{\psi'}\right]\right\| \le \rho\left\|\widehat{\Theta}(k) - \overline{\Theta}_{\psi'}\right\|, \quad 0 < \rho < 1 \quad \text{(A.6.6)}$$

with ρ independent of ψ'. Now given $\epsilon > 0$, we can find $\|\widehat{\Theta}(k) - \overline{\Theta}_{\psi'}\|$ small enough such that

$$\left\|\widehat{\Theta}(k+1) - \overline{\Theta}_{\psi'}\right\| \le \left\|\left[I - \mu \boldsymbol{H}(\overline{\Theta})\right] \cdot \left[\overline{\Theta}(k) - \overline{\Theta}_{\psi'}\right]\right\| + \epsilon\left\|\widehat{\Theta}(k) - \overline{\Theta}_{\psi'}\right\|.$$

small enough $\epsilon > 0$

$$\le (\rho + \epsilon)\left\|\widehat{\Theta}(k) - \overline{\Theta}_{\psi'}\right\| \quad \text{(A.6.8)}$$

with $\rho + \epsilon < 1$

$$\left\|\widehat{\Theta}(k+1) - \overline{\Theta}_{\psi'}\right\| < \left\|\widehat{\Theta}(k) - \overline{\Theta}_{\psi'}\right\|. \quad \text{(A.6.7)}$$

Therefore, for $\widehat{\Theta}(k)$ in the neighborhood of the manifold $\overline{\Theta}_\psi$, we have

$$V(k+1) = \inf_\psi \left\|\widehat{\Theta}(k+1) - \overline{\Theta}_\psi\right\| \le \left\|\widehat{\Theta}(k+1) - \overline{\Theta}_{\psi'}\right\| < \left\|\widehat{\Theta}(k) - \overline{\Theta}_{\psi'}\right\| = V(k). \quad \text{(A.6.9)}$$

We hereby conclude that the 1-manifold $\overline{\Theta}_\psi$ is locally stable. □

$$\boldsymbol{H}(\overline{\theta}) = \begin{bmatrix} E\{f(x_{k-n})x_k^2\} & 0 & 0 & \cdots & E\{f(x_{k-n})x_k x_{k-n}\} \\ 0 & E\{f(x_{k-n})\}E\{x_k^2\} & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ E\{f(x_{k-n})x_k x_{k-n}\} & 0 & 0 & \cdots & E\{f(x_{k-n})x_{k-n}^2\} \end{bmatrix}. \quad \text{(A.5.3)}$$

Proof of Result 4: It can easily be verified that

$$\begin{aligned} \overline{\Theta}_\psi' X_k X_k' \overline{\Theta}_\psi &= \overline{\Theta}_0' X_k X_k' \overline{\Theta}_0, \\ X_k X_k' T_\psi &= T_\psi X_k X_k', \quad \text{and} \\ T_\psi T_\psi' &= T_\psi' T_\psi = I. \end{aligned} \tag{A.7.1}$$

These identities and (4.8) may be used in (4.11) to prove the simple relationship

$$\boldsymbol{H}(\overline{\Theta}_\psi) = T_\psi \boldsymbol{H}(\overline{\Theta}_0) T_\psi'. \tag{A.7.2}$$

Thus, if λ is a positive eigenvalue and v the corresponding eigenvector of $\boldsymbol{H}(\overline{\Theta}_0)$, such that

$$\boldsymbol{H}(\overline{\Theta}_0) v = \lambda v \tag{A.7.3}$$

then

$$\boldsymbol{H}(\overline{\Theta}_\psi) T_\psi v = T_\psi \boldsymbol{H}(\overline{\Theta}_0) T_\psi' T_\psi v = T_\psi \boldsymbol{H}(\overline{\Theta}_0) v = \lambda T_\psi v. \tag{A.7.4}$$

Proving that λ is an eigenvalue and $T_\psi v$ the corresponding eigenvector of the Hessian $\boldsymbol{H}(\overline{\Theta}_\psi)$. Because T_ψ is full-rank $T_\psi v \neq 0$, and so $\boldsymbol{H}(\overline{\Theta}_\psi)$ has the same set of eigenvalues as $\boldsymbol{H}(\overline{\Theta}_0)$ for all ψ.

In addition, if the eigenvector v_0 of $\boldsymbol{H}(\overline{\Theta}_0)$ corresponding to the zero eigenvalue, is tangent to the one-dimensional manifold $\overline{\Theta}_\psi$ at $\psi = 0$, i.e.,

$$v_0 = \frac{d}{d\psi} \overline{\Theta}_\psi \bigg|_{\psi=0} \tag{A.7.5}$$

then the corresponding eigenvector $T_\psi v_0$ of $\boldsymbol{H}(\overline{\Theta}_\psi)$ clearly satisfies

$$\begin{aligned} T_\psi v_0 &= T_\psi \frac{d}{d\sigma} \overline{\Theta}_\sigma \bigg|_{\sigma=0} = T_\psi \frac{d}{d\sigma} T_\sigma \overline{\Theta}_0 \bigg|_{\sigma=0} \\ &= T_\psi \begin{bmatrix} -I\sin\sigma & I\cos\sigma \\ -I\cos\sigma & -I\sin\sigma \end{bmatrix} \overline{\Theta}_0 \bigg|_{\sigma=0} \\ &= T_\psi \begin{bmatrix} 0 & I \\ -I & 0 \end{bmatrix} \overline{\Theta}_0 = \begin{bmatrix} -I\sin\psi & I\cos\psi \\ -I\cos\psi & -I\sin\psi \end{bmatrix} \overline{\Theta}_0 \\ &= \frac{d}{d\sigma} T_\sigma \overline{\Theta}_0 \bigg|_{\sigma=\psi} = \frac{d}{d\sigma} \overline{\Theta}_\sigma \bigg|_{\sigma=\psi} \end{aligned} \tag{A.7.6}$$

which is the tangent of the 1-manifold at ψ. □

Proof of Stability Condition 2: Our assumptions are those which constrain attention to symmetric complex constellations, i.e.,

$$E\{\mathrm{Re}^2(a_k)\} = E\{\mathrm{Im}^2(a_k)\} = E|a_k|^2/2 \quad \text{and} \quad Ea_k^2 = 0. \tag{A.8.1}$$

Hence, it follows easily that $E\{\mathrm{Re}(a_k)\,\mathrm{Im}(a_k)\} = 0$, and it can be shown for the $AR(n)$ channel defined by (3.1) ($|\alpha| < 1$) that

$$\begin{aligned} E\Big\{\mathrm{Re}(x_k)^2\Big\} &= E\Big\{\mathrm{Im}(x_k)^2\Big\} = E\Big\{|x_k|^2\Big\}\Big/2 \\ &\text{and} \quad E\{\mathrm{Re}(x_k)\,\mathrm{Im}(x_k)\} = 0. \end{aligned} \tag{A.8.2}$$

From the above equations and the independency condition of x_{k-i} as well as (3.3), we have many zero entries off diagonal in the Hessian (4.13), except at entries involving x_{k-n} and x_k. Thus, through the use of all the independency relations plus (A.8.1)–(A.8.2), it can be shown as follows below in (A.8.3), in which we have let

$$\gamma = \frac{E\Big\{|x_k|^{2(p-1)}\Big\}}{E\Big\{|x_k|^{2p}\Big\}} - \frac{E\Big\{|x_k|^{p-2}\Big\}}{2E\{|x_k|^p\}}. \tag{A.8.4}$$

$$\begin{aligned} \boldsymbol{H}(\overline{\Theta}_0) = {} & 2E\Big\{\Big[(p-1)|\overline{\Theta}_0|^{2(p-2)}|x_{k-n}|^{2(p-2)} - R_p(p/2-1)|\overline{\Theta}_0|^{p-4}|x_{k-n}|^{p-4}\Big] X_k X_k' \overline{\Theta}_0 \overline{\Theta}_0' X_k X_k'\Big\} \\ & + E\Big\{\Big(|\overline{\Theta}_0|^{2(p-1)}|x_{k-n}|^{2(p-1)} - R_p|\overline{\Theta}_0|^{p-2}|x_{k-n}|^{p-2}\Big) X_k X_k'\Big\} \\ = {} & pR_p|\overline{\Theta}_0|^{p-2} E\{|x_k|^p\} \end{aligned}$$

$$\cdot \begin{bmatrix} \mathrm{Re}^2(\alpha) + E|a_k^2|\gamma & 0 & \cdots & 0 & -\mathrm{Re}(\alpha) & -\mathrm{Re}(\alpha)\,\mathrm{Im}(\alpha) & 0 & \cdots & 0 & 0 \\ 0 & E\Big\{|x_k|^2\Big\}\gamma & \cdots & 0 & 0 & 0 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & E\Big\{|x_k|^2\Big\}\gamma & 0 & 0 & 0 & \cdots & 0 & 0 \\ -\mathrm{Re}(\alpha) & 0 & \cdots & 0 & 1 & \mathrm{Im}(\alpha) & 0 & \cdots & 0 & 0 \\ -\mathrm{Re}(\alpha)\,\mathrm{Im}(\alpha) & 0 & \cdots & 0 & \mathrm{Im}(\alpha) & \mathrm{Im}^2(\alpha) + E|a_k^2|\gamma & 0 & \cdots & 0 & 0 \\ 0 & 0 & \cdots & 0 & 0 & 0 & E\Big\{|x_k|^2\Big\}\gamma & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 0 & 0 & 0 & 0 & \cdots & E\Big\{|x_k|^2\Big\}\gamma & 0 \\ 0 & 0 & \cdots & 0 & 0 & 0 & 0 & \cdots & 0 & 0 \end{bmatrix} \tag{A.8.3}$$

It can be shown that this Hessian has the following eigenvalues:

$$\lambda_1, \lambda_2 = \frac{1}{2} p R_p \left|\overline{\Theta}_0\right|^{p-2} E\{|x_k|^p\} \left[1 + |\alpha|^2 + \gamma E\left\{|a_k|^2\right\} \pm \sqrt{\left(1 + |\alpha|^2 + \gamma E\left\{|a_k|^2\right\}\right)^2 - 4\gamma E\left\{|a_k|^2\right\}} \right] \quad \text{(A.8.5a)}$$

$$\lambda_3 = p R_p \left|\overline{\Theta}_0\right|^{p-2} E\{|x_k|^p\} E\left\{|a_k|^2\right\} \gamma, \quad \text{(A.8.5b)}$$

$$\lambda_i = p R_p \left|\overline{\Theta}_0\right|^{p-2} E\{|x_k|^p\} E\left\{|x_k|^2\right\} \gamma \quad \forall\, 3 < i \le 2n+1, \quad \text{(A.8.5c)}$$

$$\lambda_{2n+2} = 0. \quad \text{(A.8.5d)}$$

Therefore, the Hessian $\boldsymbol{H}(\overline{\Theta}_0)$ is nonnegative definite and has rank $2n+1$ if and only if

$$\gamma = \frac{E\left\{|x_k|^{2(p-1)}\right\}}{E\left\{|x_k|^{2p}\right\}} - \frac{E\left\{|x_k|^{p-2}\right\}}{2E\{|x_k|^p\}} > 0 \quad \text{(A.8.6)}$$

and the eigenvector corresponding to its zero eigenvalue is $v_0 = [0\ 0\ \cdots\ 0\ 1]'$. Using the results of Lemma 4, we have thus proved Stability Condition 2. □

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] Y. Sato, "A method of self-recovering equalization for multi-level amplitude modulation," *IEEE Trans. Commun.*, vol. COM-23, pp. 679–682, June 1975.

[3] 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.

[4] J.E. Mazo, "Analysis of decision-directed equalizer convergence," *Bell Syst. Tech. J.*, pp. 1857–1876, Dec. 1980.

[5] O. Macchi and E. Eweda, "Convergence analysis of self-adaptive equalizers," *IEEE Trans. Inform. Theory*, vol. IT-30, pp. 162–176, Mar. 1984.

[6] R. Kumar, "Convergence of a decision-directed adaptive equalizer," in *Proc. Conf. Decision Contr.*, Long Beach, CA, 1983, vol. 3, pp. 1319–1324.

[7] G.J. Foschini, "Equalization without altering or detecting data," *AT&T Tech. J.*, vol. 64, pp. 1885–1911, Oct. 1985.

[8] J.R. Treichler and M.G. Larimore, "New processing techniques based on the constant modulus adaptive algorithm," *IEEE Trans. Acoust., Speech, Signal Processing,* vol. ASSP-33, pp. 420–431, April 1985.

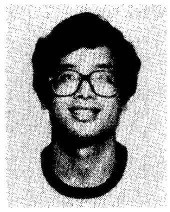
[9] J.R. Treichler and M.G. Agee, "A new approach to multipath correction of constant modulus signals," *IEEE Trans. Acoust., Speech, Signal Processing,* vol. ASSP-28, pp. 349–472, Apr. 1983.

[10] C.R. Johnson, Jr., S. Dasgupta, and W.A. Sethares, "Averaging analysis of local stability of a real constant modulus algorithm adaptive filter," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 36, pp. 900–910, June 1988.

[11] B.D.O. Anderson, R.R. Bitmead, C.R. Johnson, Jr., P.V. Kokotovic, R.L. Kosut, I.M.Y. Mareels, L. Praly, and B.D. Riedle, *Stability of Adaptive Systems: Passivity and Averaging Analysis.* Cambridge, MA: M.I.T. Press, 1986.

[12] S. Verdú, "On the selection of memoryless adaptive laws for blind equalization in binary communications," in *Proc. 6th Int. Conf. Anal. Optimiz. Syst.*, Nice, France, June 1984, pp. 239–249.

[13] G.J. Rey, "Local stability analysis of the constant modulus algorithm for adaptive equalization," M.S. thesis, Cornell Univ., Ithaca, NY, Jan. 1989.



Zhi Ding (S'88–M'90) was born in Harbin, China, in 1962. He received the B.Eng. degree from the Department of Radio Engineering, Nanjing Institute of Technology, Nanjing, China, in July 1982, the M.A.Sc. degree from the Department of Electrical Engineering, University of Toronto, Toronto, Ont., Canada, in May 1987, and 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 currently an Assistant Professor in the Department of Electrical Engineering. His research interests include adaptive systems theory, data communications, and signal processing.

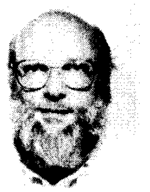
Rodney A. Kennedy (S'86–M'88) was born in Sydney, Australia, on October 23, 1960. He received the B.E.(Hons.) degree in electrical engineering from the University of New South Wales, NSW, Australia, in 1982, the M.E. degree in digital control theory from the University of Newcastle, Australia, in 1986, and the Ph.D. degree from the Department of Systems Engineering, Australia National University, Canberra, Australia, in 1988.

From 1983 to 1986 he worked on antenna design problems for the C.S.I.R.O. Division of Radiophysics, Sydney, Australia. Currently he is a Research Fellow in the Department of Systems Engineering, Australia National University, Canberra. His research activities are in digital communications, signal processing, and stochastic processes.

Brian D. O. Anderson (S'62–M'66–SM'74–F'75) was born in Sydney, Australia. He received the undergraduate degrees in pure mathematics and electrical engineering from the University of Sydney, Sydney, Australia, and the Ph.D. degree in electrical engineering from Stanford University, Stanford, CA.

Following completion of his education, he worked in industry in Silicon Valley and served as an Assistant Professor in the Department of Electrical Engineering at Stanford. In 1967, he became Foundation Professor of Electrical Engineering at the University of Newcastle, New South Wales, Australia, where he remained until the end of 1981. At that time, he became Professor of Systems Engineering at the Australian National University, Canberra. His research interests are in control engineering and signal processing, and he has coauthored a number of books in these fields, as well as research papers.

Dr. Anderson is a Fellow of the Royal Society, the Australian Academy of Science, the Australian Academy of Technological Sciences and Engineering, and an Honorary Fellow of the Institution of Engineers, Australia. He is serving a term as President of the International Federation of Automatic Control from 1990 to 1993.



C. Richard Johnson, Jr. (S'74–M'77–SM'82–F'89) was born in Macon, GA, in 1950. He received the Ph.D. degree in electrical engineering with minors in engineering-economic systems and art history from Stanford University, Stanford, CA, in 1977.

He is currently a Professor and Associate Director of the School of Electrical Engineering at Cornell University, Ithaca, NY. His research interests are in adaptive parameter estimation theory as applied to digital filtering, communications, identification, and control.

Dr. Johnson has served as an Editor of the *International Journal of Adaptive Control and Signal Processing;* and as an Associate Editor for the IEEE TRANSACTIONS ON ACOUSTICS, SPEECH, AND SIGNAL PROCESSING; the IEEE TRANSACTIONS ON AUTOMATIC CONTROL; *Automatica*; and *Systems and Control Letters.*