Zhi Ding. “Adaptive Filters for Blind Equalization.”
2000 CRC Press LLC. <http://www.engnetbase.com>.

Adaptive Filters for Blind Equalization

Zhi Ding
Auburn University

24.1 Introduction

One of the earliest and most successful applications of adaptive filters is adaptive channel equalization in digital communication systems. Using the standard least mean LMS algorithm, an adaptive equalizer is a finite-impulse-response FIR filter whose desired reference signal is a known training sequence sent by the transmitter over the unknown channel. The reliance of an adaptive channel equalizer on a training sequence requires that the transmitter cooperates by (often periodically) resending the training sequence, lowering the effective data rate of the communication link.

In many high-data-rate bandlimited digital communication systems, the transmission of a training sequence is either impractical or very costly in terms of data throughput. Conventional LMS adaptive filters depending on the use of training sequences cannot be used. For this reason, blind adaptive channel equalization algorithms that do not rely on training signals have been developed. Using these "blind" algorithms, individual receivers can begin self-adaptation without transmitter assistance. This ability of blind startup also enables a blind equalizer to self-recover from system breakdowns. This self-recovery ability is critical in broadcast and multicast systems where channel variation often occurs.

In this section, we provide an introduction to the basics of blind adaptive equalization. We describe commonly used blind algorithms, highlight important issues regarding convergence properties of various blind equalizers, outline common initialization tactics, present several open problems, and discuss recent advances in this field.

24.2 Channel Equalization in QAM Data Communication Systems

In data communication, digital signals are transmitted by the sender through an analog channel to the receiver. Nonideal analog media such as telephone cables and radio channels typically distort the transmitted signal.

The problem of blind channel equalization can be described using the simple system diagram shown in Fig. 24.1. The complex baseband model for a typical QAM (quadrature amplitude modulated) data communication system consists of an unknown linear time-invariant (LTI) channel $h(t)$ which represents all the interconnections between the transmitter and the receiver at baseband. The matched filter is also included in the LTI channel model. The baseband-equivalent transmitter generates a sequence of complex-valued random input data $\{a(n)\}$, each element of which belongs to a complex alphabet $\mathcal{A}$ (or constellation) of QAM symbols. The data sequence $\{a(n)\}$ is sent through a baseband-equivalent complex LTI channel whose output $x(t)$ is observed by the receiver. The function of the receiver is to estimate the original data $\{a(n)\}$ from the received signal $x(t)$.

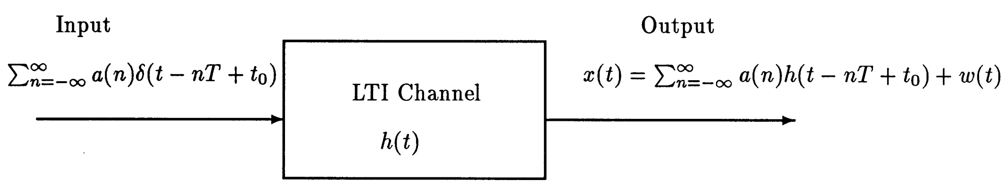

FIGURE 24.1: Baseband representation of a QAM data communication system.

For a causal and complex-valued LTI communication channel with impulse response $h(t)$, the input/output relationship of the QAM system can be written as

$$x(t) = \sum_{n=-\infty}^{\infty} a(n)h(t-nT+t_0) + w(t), \quad a(n) \in \mathcal{A}\,, \tag{24.1}$$

where T is the symbol (or baud) period. Typically the channel noise $w(t)$ is assumed to be stationary, Gaussian, and independent of the channel input $a(n)$.

In typical communication systems, the matched filter output of the channel is sampled at the known symbol rate $1/T$ assuming perfect timing recovery. For our model, the sampled channel output

$$x(nT) = \sum_{k=-\infty}^{\infty} a(k)h(nT-kT+t_0) + w(nT) \tag{24.2}$$

is a discrete time stationary process. Equation (24.2) relates the channel input to the sampled matched filter output. Using the notations

$$x(n) \triangleq x(nT), \qquad w(n) \triangleq w(nT), \qquad \text{and} \quad h(n) \triangleq h(nT+t_0)\,, \tag{24.3}$$

the relationship in (24.2) can be written as

$$x(n) = \sum_{k=-\infty}^{\infty} a(k)h(n-k) + w(n) \,. \tag{24.4}$$

When the channel is nonideal, its impulse response $h(n)$ is nonzero for $n \neq 0$. Consequently, undesirable signal distortion is introduced as the channel output $x(n)$ depends on multiple symbols in $\{a(n)\}$. This phenomenon, known as *intersymbol* interference (ISI), can severely corrupt the transmitted signal. ISI is usually caused by limited channel bandwidth, multipath, and channel fading in digital communication systems. A simple memoryless decision device acting on $x(n)$ may not be able to recover the original data sequence under strong ISI. Channel equalization has proven to be an effective means of significant ISI removal. A comprehensive tutorial on nonblind adaptive channel equalization by Qureshi [2] contains detailed discussions on various aspects of channel equalization.

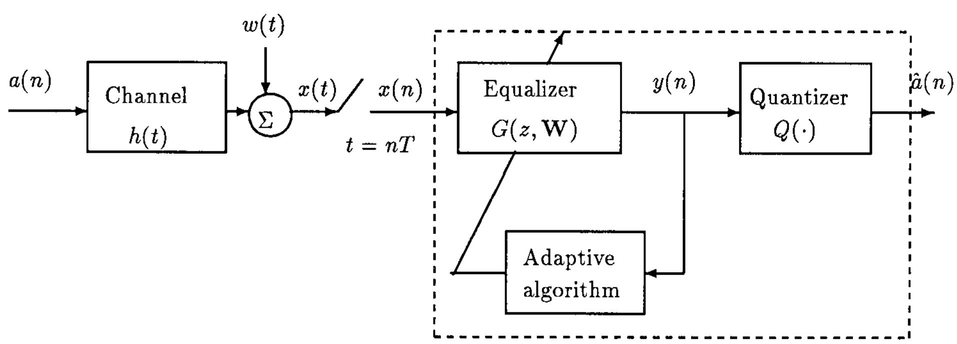

FIGURE 24.2: Adaptive blind equalization system.

Figure 24.2 shows the combined communication system with adaptive equalization. In this system, the equalizer $G(z, \mathbf{W})$ is a linear FIR filter with parameter vector $\mathbf{W}$ designed to remove the distortion caused by channel ISI. The goal of the equalizer is to generate an output signal $y(n)$ that can be quantized to yield a reliable estimate of the channel input data as

$$\hat{a}(n) = Q\left(y(n)\right) = a(n-\delta) \,, \tag{24.5}$$

where δ is a constant integer delay. Typically any constant but finite amount of delay introduced by the combined channel and equalizer is acceptable in communication systems.

The basic task of equalizing a linear channel can be translated to that task of identifying the equivalent discrete channel, defined in z-transform notation as

$$H(z) = \sum_{k=0}^{\infty} h(k) z^{-k} \,. \tag{24.6}$$

With this notation, the channel output becomes

$$x(n) = H(z)a(n) + w(n) \tag{24.7}$$

where $H(z)a(n)$ denotes linear filtering of the sequence $a(n)$ by the channel and $w(n)$ is a white (for a root-raised-cosine matched filter [2]) stationary noise with constant power spectrum N_0. Once

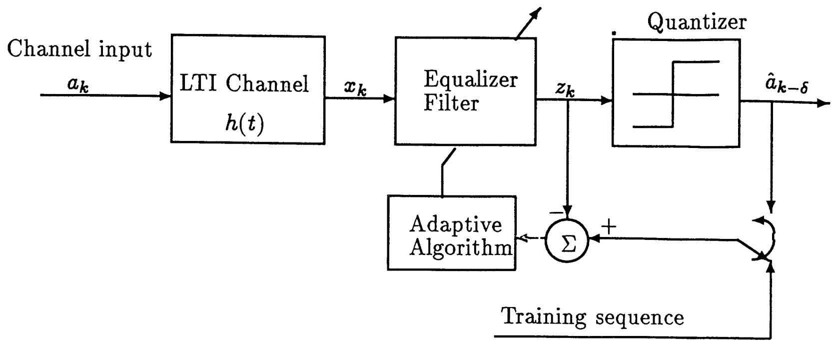

FIGURE 24.3: Decision-directed channel equalization algorithm.

the channel has been identified, the equalizer can be constructed according to the minimum mean square error (MMSE) criterion between the desired signal $a(n-\delta)$ and the output $y(n)$ as

$$G_{mmse}(z, \mathbf{W}) = \frac{H^*(z^{-1})z^{-\delta}}{H(z)H^*(z^{-1}) + N_0}\,, \tag{24.8}$$

where $*$ denotes complex conjugate. Alternatively, if the zero-forcing (ZF) criterion is employed, then the optimum ZF equalizer is

$$G_{zf}(z, \mathbf{W}) = \frac{z^{-\delta}}{H(z)}\,, \tag{24.9}$$

which causes the combined channel-equalizer response to become a purely δ-sample delay with zero ISI. ZF equalizers tend to perform poorly when the channel noise is significant and when the channels $H(z)$ have zeros near the unit circle.

Both the MMSE equalizer (24.8) and the ZF equalizer (24.9) are of a general infinite impulse response (IIR) form. However, adaptive linear equalizers are usually implemented as FIR filters due to the difficulties inherent in adapting IIR filters. Adaptation is then based on a well-defined criterion such as the MMSE between the ideal IIR and truncated FIR impulse responses or the MMSE between the training signal and the equalizer output.

24.3 Decision-Directed Adaptive Channel Equalizer

Adaptive channel equalization was first developed by Lucky [1] for telephone channels. Figure 24.3 depicts the traditional adaptive equalizer. The equalizer begins adaptation with the assistance of a known training sequence initially transmitted over the channel. Since the training signal is known, standard gradient-based adaptive algorithms such as the LMS algorithm can be used to adjust the equalizer coefficients to minimize the mean square error (MSE) between the equalizer output and the training sequence. It is assumed that the equalizer coefficients are sufficiently close to their optimum values and that much of the ISI is removed by the end of the training period. Once the channel input sequence $\{a(n)\}$ can be accurately recovered from the equalizer output through a memoryless decision device such as a quantizer, the system is switched to the decision-directed mode whereby the adaptive equalizer obtains its reference signal from the decision output.

One can construct a blind equalizer by employing decision-directed adaptation without a training sequence. The algorithm minimizes the MSE between the quantizer output

$$\hat{a}(n-\delta) = Q(y(n)) \tag{24.10}$$

and the equalizer output $y(n)$. Naturally, the performance of the decision-directed algorithm depends on the accuracy of the estimate $Q(y(n))$ for the true symbol $a(n-\delta)$. Undesirable convergence to a local minimum with severe residual ISI can occur in this situation such that $Q(y(n))$ and $a(n-\delta)$ differ sufficiently often. Thus, the challenge of blind equalization lies in the design of special adaptive algorithms that eliminate the need for training without compromising the desired convergence to near the optimum MMSE or ZF equalizer coefficients.

24.4 Basic Facts on Blind Adaptive Equalization

In blind equalization, the desired signal or input to the channel is unknown to the receiver, except for its probabilistic or statistical properties over some known alphabet $\mathcal{A}$. As both the channel $h(n)$ and its input $a(n)$ are unknown, the objective of blind equalization is to recover the unknown input sequence based solely on its probabilistic and statistical properties.

The first comprehensive analytical study of the blind equalization problem was presented by Benveniste, Goursat, and Ruget in 1980 [3]. In fact, the very term "blind equalization" can be attributed to Benveniste and Goursat from the title of their 1984 paper [4]. The seminal paper of Benveniste et al. [3] established the connection between the task of blind equalization and the use of higher order statistics of the channel output. Through rigorous analysis, they generalized the original Sato algorithm [5] into a class of algorithms based on non-MSE cost functions. More importantly, the convergence properties of the proposed algorithms were carefully investigated. Based on the work of [3], the following facts about blind equalization are generally noted:

1. Second order statistics of $x(n)$ alone only provide the magnitude information of the linear channel and are insufficient for blind equalization of a mixed phase channel $H(z)$ containing zeros inside and outside the unit circle in the z-plane.
2. A mixed phase linear channel $H(z)$ cannot be identified from its outputs when the input signal is i.i.d. Gaussian, since only second order statistical information is available.
3. Although the exact inverse of a nonminimum phase channel is unstable, a truncated anticausal expansion can be delayed by δ to allow a causal approximation to a ZF equalizer.
4. ZF equalizers cannot be implemented for channels $H(z)$ with zeros on the unit circle.
5. The symmetry of QAM constellations $\mathcal{A} \subset \mathbb{C}$ causes an inherent phase ambiguity in the estimate of the channel input sequence or the unknown channel when input to the channel is uniformly distributed over $\mathcal{A}$. This phase ambiguity can be overcome by differential encoding of the channel input.

Due to the absence of a training signal, it is important to exploit various available information about the input symbol and the channel output to improve the quality of blind equalization. Usually, the following information is available to the receiver for blind equalization:

- The power spectral density (PSD) of the channel output signal $x(t)$, which contains information on the magnitude of the channel transfer function;
- The higher-order statistics (HOS) of the T-sampled channel output $\{x(kT)\}$, which contains information on the phase of the channel transfer function;
- Cyclostationary second and higher order statistics of the channel output signal $x(t)$, which contain additional phase information of the channel; and
- The finite channel input alphabet, which can be used to design quantizers or decision devices with memory to improve the reliability of the channel input estimate.

Naturally in some cases, these information sources are not necessarily independent as they contain overlapping information. Efficient and effective blind equalization schemes are more likely to be designed when all useful information is exploited at the receiver. We now describe various algorithms for blind channel identification and equalization.

24.5 Adaptive Algorithms and Notations

There are basically two different approaches to the problem of blind equalization. The stochastic gradient descent (SGD) approach iteratively minimizes a chosen cost function over all possible choices of the equalizer coefficients, while the statistical approach uses sufficient stationary statistics collected over a block of received data for channel identification or equalization. The latter approach often exploits higher order or cyclostationary statistical information directly. In this discussion, we focus on the the adaptive online equalization methods employing the the gradient descent approach, as these methods are most closely related to other topics in this chapter. Consequently, the design of special, non-MSE cost functions that implicitly exploits the HOS of the channel output is the key issue in our methods and discussions.

For reasons of practicality and ease of adaptation, a linear channel equalizer is typically implemented as an FIR filter $G(z,\ \mathbf{W})$. Denote the equalizer parameter vector as

$$\mathbf{W} \stackrel{\Delta}{=} [w_0 \ \ w_1 \ \ \cdots \ \ w_m\,]^T\,, \quad m < \infty\,.$$

In addition, define the received signal vector as

$$\mathbf{X}(n) \stackrel{\Delta}{=} [x(n) \ \ x(n-1) \ \ \ldots \ \ x(n-m)\,]^T\,. \tag{24.11}$$

The output signal of the linear equalizer is thus

$$\begin{aligned} y(n) &= \mathbf{W}^T\mathbf{X}(n) \\ &= G(z,\mathbf{W})\{x(n)\}\,, \end{aligned} \tag{24.12}$$

where we have defined the equalizer transfer function as

$$G(z,\ \mathbf{W}) = \sum_{i=0}^{m} w_i z^{-i}\,. \tag{24.13}$$

All the ISI is removed by a ZF equalizer if

$$H(z)G(z,\mathbf{W}) = gz^{-\delta}, \quad g \neq 0 \tag{24.14}$$

such that the noiseless equalizer output becomes $y(n) = ga(n-\delta)$, where g is a complex-valued scaling factor. Hence, a ZF equalizer attempts to achieve the inverse of the channel transfer function with a possible gain difference g and/or a constant time delay δ.

Denoting the parameter vector of the equalizer at sample instant n as $\mathbf{W}(n)$, the conventional LMS adaptive equalizer employing a training sequence is given by

$$\mathbf{W}(n+1) = \mathbf{W}(n) + \mu[a(n-\delta) - y(n)]\mathbf{X}(n)\,, \tag{24.15}$$

where $\cdot^*$ denotes complex conjugates and μ is a small positive stepsize. Naturally, this algorithm requires that the channel input $a(n-\delta)$ be available. The equalizer iteratively minimizes the MSE cost function

$$E\left\{|e_n|^2\right\} = E\{|a(n-\delta) - y(n)|^2\}\,.$$

If the MSE is so small after training that the equalizer output $y(n)$ is a close estimate of the true channel input $a(n-\delta)$, then $Q(y(n))$ can replace $a(n-\delta)$ in a decision-directed algorithm that continues to track modest time-variations in the channel dynamics [2].

In blind equalization, the channel input $a(n-\delta)$ is unavailable, and thus different minimization criteria are explored. The crudest blind equalization algorithm is the decision-directed scheme that updates the adaptive equalizer coefficients as

$$\mathbf{W}(n+1) = \mathbf{W}(n) + \mu[Q(y(n)) - y(n)]\mathbf{X}^*(n) . \tag{24.16}$$

The performance of the decision-directed algorithm depends on how close $\mathbf{W}(n)$ is to its optimum setting $\mathbf{W}_{opt}$ under the MMSE or the ZF criterion. The closer $\mathbf{W}(n)$ is to $\mathbf{W}_{opt}$, the smaller the ISI is and the more accurate the estimate $Q(y(n))$ is to $a(n-\delta)$. Consequently, the algorithm in (24.16) is likely to converge to $\mathbf{W}_{opt}$ if $\mathbf{W}(n)$ is initially close to $\mathbf{W}_{opt}$. The validity of this intuitive argument is shown analytically in [6, 7]. On the other hand, $\mathbf{W}(n)$ can also converge to parameter values that do not remove sufficient ISI from certain initial parameter values $\mathbf{W}(0)$, as $Q(y(n)) \neq a(n-\delta)$ sufficiently often in some cases [6, 7].

The ability of the equalizer to achieve the desired convergence result when it is initialized with sufficiently small ISI accounts for the key role that the decision-directed algorithm plays in channel equalization. In the system of Fig. 24.3, the training session is designed to help $\mathbf{W}(n)$ converge to a parameter vector such that most of the ISI has been removed, from which adaptation can be switched to the decision-directed mode. Without direct training, a blind equalization algorithm is therefore used to provide a good initialization for the decision-directed equalizer because of the decision-directed equalizer's poor convergence behavior under high ISI.

24.6 Mean Cost Functions and Associated Algorithms

Under the zero-forcing criterion, the objective of the blind equalizer is to adjust $\mathbf{W}(n)$ such that (24.14) can be achieved using a suitable rule of self-adaptation. We now describe the general methodology of blind adaptation and introduce several popular algorithms.

Unless otherwise stated, we focus on the blind equalization of pulse-amplitude modulation (PAM) signals, in which the input symbol is uniformly distributed over the following M levels,

$$\{\pm(M-1)d,\ \pm(M-3)d,\ \ldots,\ \pm 3d,\ \pm d\}, \qquad M \text{ even} . \tag{24.17}$$

We study this particular case because (1) algorithms are often defined only for real signals when first developed [3, 5], and (2) the extension to complex (QAM) systems is generally straightforward [4].

Blind adaptive equalization algorithms are often designed by minimizing special non-MSE cost functions that do not involve the use of the original input $a(n)$ but still reflect the current level of ISI in the equalizer output. Define the *mean cost function* as

$$J(\mathbf{W}) \triangleq E\{\Psi(y(n))\} , \tag{24.18}$$

where $\Psi(\cdot)$ is a scalar function of its argument. The mean cost function $J(\mathbf{W})$ should be specified such that its minimum point $\mathbf{W}$ corresponds to a minimum ISI or MSE condition. Because of the symmetric distribution of $a(n)$ over $\mathcal{A}$ in 24.17, the function Ψ should be even ($\Psi(-x) = \Psi(x)$), so that both $y(n) = a(n-\delta)$ and $y(n) = -a(n-\delta)$ are desired objectives or global minima of the mean cost function.

Using 24.18, the stochastic gradient descent minimization algorithm is easily derived as [3]

$$\begin{aligned} \mathbf{W}(n+1) &= \mathbf{W}(n) - \mu \cdot \frac{\partial}{\partial \mathbf{W}(n)} \Psi(y(n)) \\ &= \mathbf{W}(n) - \mu \cdot \Psi'\left(\mathbf{X}^T(n)\mathbf{W}(n)\right)\mathbf{X}^*(n) . \end{aligned} \tag{24.19}$$

Define the first derivative of Ψ as

$$\psi(x) \triangleq \Psi'(x) = \frac{\partial}{\partial x}\Psi(x) .$$

The resulting blind equalization algorithm can then be written as

$$\mathbf{W}(n+1) = \mathbf{W}(n) - \mu\psi(\mathbf{X}^T(n)\mathbf{W}(n))\mathbf{X}^*(n) . \tag{24.20}$$

Hence, a blind equalizer can either be defined by its cost function $\Psi(x)$, or equivalently, by the derivative $\psi(x)$ of its cost function, which is also called the *error function* since it replaces the prediction error in the LMS algorithm. Correspondingly, we have the following relationship:

$$\text{Minima of the mean cost } J(\mathbf{W}) \iff \text{Stable equilibria of the algorithm in (24.20)} .$$

The design of the blind equalizer thus translates into the selection of the function Ψ (or ψ) such that local minima of $J(\mathbf{W})$, or equivalently, the locally stable equilibria of the algorithm (24.20) correspond to a significant removal of ISI in the equalizer output.

24.6.1 The Sato Algorithm

The first blind equalizer for multilevel PAM signals was introduced by Sato [5] and is defined by the error function

$$\psi_1(y(n)) = y(n) - R_1 \operatorname{sgn}(y(n)) , \tag{24.21}$$

where

$$R_1 \triangleq \frac{E|a(n)|^2}{E|a(n)|} .$$

Clearly, the Sato algorithm effectively replaces $a(n-\delta)$ with $R_1\operatorname{sgn}(y(n))$, known as the slicer output. The multilevel PAM signal is viewed as an equivalent binary input signal in this case, so that the error function often has the same sign for adaptation as the LMS error $y(n) - a(n-\delta)$.

24.6.2 BGR Extensions of Sato Algorithm

The Sato algorithm was extended by Benveniste, Goursat, and Ruget [3] who introduced a class of error functions given by

$$\psi_b(y(n)) = \tilde{\psi}(y(n)) - R_b \operatorname{sgn}(y(n)) , \tag{24.22}$$

where

$$R_b \triangleq \frac{E\{\tilde{\psi}(a(n))a(n)\}}{E|a(n)|} . \tag{24.23}$$

Here, $\tilde{\psi}(x)$ is an odd and twice differentiable function satisfying

$$\tilde{\psi}''(x) \geq 0, \quad \forall x \geq 0 . \tag{24.24}$$

The use of the function $\tilde{\psi}$ generalizes the linear function $\tilde{\psi}(x) = x$ in the Sato algorithm. The class of algorithms satisfying (24.22) and (24.24) are called *BGR algorithms*. They are individually represented by the explicit specification of the $\tilde{\psi}$ function, as with the Sato algorithm.

The generalization of these algorithms to complex signals (QAM) and complex equalizer parameters is straightforward by separating signals into their real and the imaginary as

$$\psi_b(y(n)) = \tilde{\psi}(Re[y(n)]) - R_b\operatorname{sgn}(Re[y(n)]) + j\{\tilde{\psi}(Im[y(n)]) - R_b\operatorname{sgn}(Im[y(n)])\} . \tag{24.25}$$

24.6.3 Constant Modulus or Godard Algorithms

Integrating the Sato error function $\psi_1(x)$ shows that the Sato algorithm has an equivalent cost function

$$\Psi_1(y(n)) = \frac{1}{2}\left(|y(n)| - R_1\right)^2 .$$

This cost function was generalized by Godard into another class of algorithms that are specified by the cost functions [8]

$$\Psi_q(y(n)) = \frac{1}{2q}\left(|y(n)|^q - R_q\right)^2 , \quad q = 1, 2, \ldots \tag{24.26}$$

where

$$R_q \triangleq \frac{E|a(n)|^{2q}}{E|a(n)|^q} .$$

This class of *Godard algorithms* is indexed by the positive integer q. Using the stochastic gradient descent approach, the Godard algorithms given by

$$\mathbf{W}(n+1) = \mathbf{W}(n) - \mu(|\mathbf{X}(n)^H\mathbf{W}(n)|^q - R_q)|\mathbf{X}(n)^T\mathbf{W}(n)|^{q-2}\mathbf{X}(n)^T\mathbf{W}(n)\mathbf{X}^*(n) . \tag{24.27}$$

The Godard algorithm for the case $q = 2$ was independently developed as the "constant modulus algorithm" (CMA) by Treichler and co-workers [9] using the philosophy of property restoral. For channel input signal that has a constant modulus $|a(n)|^2 = R_2$, the CMA equalizer penalizes output samples $y(n)$ that do not have the desired constant modulus characteristics. The modulus error is simply

$$e(n) = |y(n)|^2 - R_2 ,$$

and the squaring of this error yields the constant modulus cost function that is the identical to the Godard cost function.

This modulus restoral concept has a particular advantage in that it allows the equalizer to be adapted independent of carrier recovery. A carrier frequency offset of Δ_f causes a possible phase rotation of the equalizer output so that

$$y(n) = |y(n)| \exp[j(2\pi\Delta_f n + \phi(n))] .$$

Because the CMA cost function is insensitive to the phase of $y(n)$, the equalizer parameter adaptation can occur independently and simultaneously with the operation of the carrier recovery system. This property also allows CMA to be applied to analog modulation signals with constant amplitude such as those using frequency or phase modulation [9].

24.6.4 Stop-and-Go Algorithms

Given the standard form of the blind equalization algorithm in (24.20), it is apparent that the convergence characteristics of these algorithms are largely determined by the sign of the error signal $\psi(y(n))$. In order for the coefficients of a blind equalizer to converge to the vicinity of the optimum MMSE solution as observed through LMS adaptation, the sign of its error signal should agree with the sign of the LMS prediction error $y(n) - a(n-\delta)$ most of the time. Slow convergence or convergence of the parameters to local minima of the cost function $J(\mathbf{W})$ that do not provide proper equalization can occur if the signs of these two errors differ sufficiently often. In order to improve the convergence properties of blind equalizers, the so-called "stop-and-go" methodology was proposed by Picchi et al. [10]. We now describe its simple concept.

The idea behind the stop-and-go algorithms is to allow adaptation “to go” only when the error function is more likely to have the correct sign for the gradient descent direction. Since there are several criteria for blind equalization, one can expect a more accurate descent direction when more than one of the existing algorithms provide the same sign of the error function. When the error signs differ for a particular output sample, parameter adaptation is “stopped”. Consider two algorithms with error functions $\psi_1(y)$ and $\psi_2(y)$. We can devise the following *stop-and-go* algorithm:

$$\mathbf{W}(k+1) = \begin{cases} \mathbf{W}(k) - \mu\psi_1(y(n))\mathbf{X}^*(n), & \text{if } \operatorname{sgn}[\psi_1(y(n))] = \operatorname{sgn}[\psi_2(y(n))]\,; \\ \mathbf{W}(k), & \text{if } \operatorname{sgn}[\psi_1(y(n))] \neq \operatorname{sgn}[\psi_2(y(n))]\,. \end{cases} \tag{24.28}$$

In their work, Picchi and Prati combined only the Sato and the decision-directed algorithms with faster convergence results through the corresponding error function

$$\psi(y(n)) = \frac{1}{2}\left(y(n) - Q(y(n))\right) + \frac{1}{2}\left|y(n) - Q(y(n))\right| \operatorname{sgn}\left(y(n) - R_1 \operatorname{sgn}(y(n))\right)\,.$$

However, given the number of existing algorithms, the stop-and-go methodology can include many different combinations of error functions. One that combines Sato and Godard algorithms was tested by Hatzinakos [11].

24.6.5 Shalvi and Weinstein Algorithms

Unlike previously introduced algorithms, the methods of Shalvi-Weinstein [12] are based on higher order statistics of the equalizer output. Define the kurtosis of the equalizer output signal $y(n)$ as

$$K_y \triangleq E\left|y(n)^4\right| - 2E^2\left|y(n)^2\right| - \left|E\left\{y(n)^2\right\}\right|^2\,. \tag{24.29}$$

The Shalvi-Weinstein algorithm maximizes $|K_y|$ subject to the constant power constraint $E|y(n)|^2 = E|a(n)^2|$. Define c_n as the combined channel-equalizer impulse response given by

$$c_n = \sum_{k=0}^{m} h_k w_{n-k}, \qquad -\infty < n < \infty\,. \tag{24.30}$$

Using the fact that $a(n)$ is i.i.d., it can be shown [13] that

$$E\left|y(n)^2\right| = E\left|a(n)^2\right| \sum_{i=-\infty}^{\infty} |c_i|^2 \tag{24.31}$$

$$K_y = K_a \sum |c_n|^4\,, \tag{24.32}$$

where K_a is the kurtosis of the channel input, a quantity that is nonzero for most QAM and PAM signals. Hence, the Shalvi-Weinstein equalizer is equivalent to the following criterion:

$$\text{maximize} \quad \sum_{n=-\infty}^{\infty} |c_n|^4 \qquad \text{subject to} \qquad \sum_{n=-\infty}^{\infty} |c_n|^2 = 1\,. \tag{24.33}$$

It can be shown [14] that there is a one-to-one correspondence between the minima of the cost function surface searched by this algorithm and those of the Godard algorithm with $q = 2$. However, the methods of adaptation given in [12] can exhibit convergence characteristics different from those of the CMA.

24.6.6 Summary

Over the years, there have been many attempts to derive new algorithms and equalization methods that are more reliable and faster than the existing methods. Nonetheless, the algorithms presented above are still the most commonly used methods in blind equalization due to their computational simplicity and practical effectiveness. In particular, CMA has proven to be useful not only in blind equalization but also in blind array signal processing systems. Because it does not rely on the accuracy of the decision device output nor the knowledge of the channel input signal constellation, CMA is a versatile algorithm that can be used not only for digital communication signals but also for analog signals that do not conform to a finite constellation alphabet.

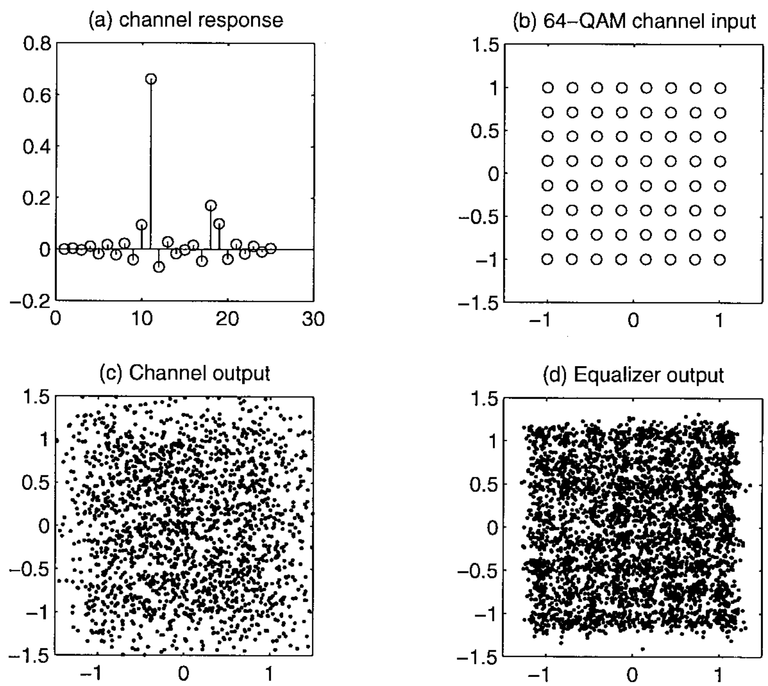

FIGURE 24.4: An example of a CMA equalizer in a 64 QAM communication system.

As a practical example, consider a QAM system in which the channel impulse response is shown in Fig. 24.4(a). This sampled composite channel response results from a continuous time system in which the transmitter and receiver filters both have identical root-raised cosine frequency response with the roll-off factor of 0.13, while the channel between the two filters is nonideal with several nondominant multipaths. The channel input signal is generated from a rectangular 64-QAM constellation as shown in Fig. 24.4(b). The channel output points are shown in Fig. 24.4(c). The channel output signal clearly has significant ISI such that a simple quantizer based on the nearest neighbor principle is likely to make many decision errors. We use a CMA equalizer with 25 parameter taps. The equalizer input is normalized by its power and a stepsize $\mu = 10^{-3}$ is used in the CMA adaptation. After 20,000 iterations, the final equalizer output after parameter convergence is shown in Fig. 24.4(d). The tighter clustering of the equalizer output shows that the decision error rate will be very low so that the equalizer can be switched to the decision-directed or decision-feedback algorithm mode at this point.

24.7 Initialization and Convergence of Blind Equalizers

The success and effectiveness of a QAM blind equalization algorithm clearly hinges on its convergence behavior in practical QAM systems with distortive channels. A desired globally convergent algorithm should only produce stable equilibria that are close to the optimum MMSE or ZF equalizer coefficients. If an equalization algorithm has local equilibria, then the initial equalizer parameter values are critical in determining the final values of parameters at convergence. Due to the analytical difficulty in locating and characterizing these local minima, most analytical studies of blind equalizers focus on the noiseless environment. For noiseless channels, the optimum MMSE and ZF equalizers are identical. The goal in the noiseless system is to remove sufficient ISI so that the open eye condition or errorless decision output, given by

$$Q(y(n)) = a(n-\delta) \, ,$$

holds.

Although the problem of blind equalization has been studied for over two decades, useful convergence analyses of most blind adaptive algorithms have proven to be difficult to perform. While some recent analytical results have helped to characterize the behavior of several popular algorithms, the overall knowledge of the behaviors of most known effective algorithms is still quite limited. Consequently, practical implementations of blind equalizers still employ heuristic measures to improve their convergence characteristics. We summarize several issues regarding the convergence and initialization of blind equalizers in this subsection.

24.7.1 A Common Analysis Approach

Although many readers may be surprised by the apparent lack of convergence proofs for most blind equalization algorithms, a closer look at the cost functions for these algorithms shows the analytical difficulty of the problem. Specifically, the stable stationary points of the blind algorithm in (24.20) correspond to the local minima of the mean cost function

$$J(\mathbf{W}) = E\left\{\Psi\left(\sum_{i=0}^{m} w_i x(n-i)\right)\right\} \, . \tag{24.34}$$

The convergence of the adaptive algorithm is thus determined by the geometry of the error function $J(\mathbf{W})$ over the equalizer parameters $\{w_i\}$. An analysis of the convergence of the algorithm in terms of its parameters $\{w_i\}$ is difficult because the statistical characterization of the channel output signal $x(n)$ is highly dependent on the channel impulse response. For this reason, most blind equalization algorithms have initially been presented with only simulation results and without a rigorous convergence analysis.

Faced with this difficulty, several researchers have studied the global behavior of the equalizer in the combined parameter space c_i of (24.30) since

$$J(\mathbf{W}) = E\left\{\Psi\left(\sum_{i=-\infty}^{\infty} w_i x(n-i)\right)\right\} = E\left\{\Psi\left(\sum_{i=-\infty}^{\infty} c_i a(n-i)\right)\right\} \, . \tag{24.35}$$

Because the probabilistic information of the signal $a(n)$ is completely known, the convergence analysis of the c_i parameters tends to be much simpler than that of the equalizer parameters w_i. The following convergence results are known from these analyses:

- For channel input signals with uniform or sub-Gaussian probability distributions, Sato and BGR algorithms are globally convergent under zero channel noise. The corresponding cost functions only have global minima at parameter settings that result in zero ISI [3].

- For uniform and discrete PAM channel input distributions, undesirable local minima of the Sato and the BGR algorithms exist that do not satisfy the open-eye condition [6, 7, 16].
- For uniform and discrete PAM (or QAM) channel input distributions, the Godard algorithm with $q = 2$ (CMA) and the Shalvi-Weinstein algorithm have no local minima under zero channel noise. Only global minima exist at parameter settings that result in zero ISI [12, 15]. In other words, all minima satisfy the zero-forcing condition

$$c_n^2 = \begin{cases} 1, & n = \delta \\ 0, & n \neq \delta \, . \end{cases} \tag{24.36}$$

24.7.2 Local Convergence of Blind Equalizers

In order for the convergence analysis of the c_i parameters to be valid for the w_i parameters, a one-to-one linear mapping must exist between the two parameter spaces. A cost function of two variables will still have the same number of minima, maxima, and saddle points after a linear one-to-one coordinate change. On the other hand, a mapping that is not one-to-one can turn a nonstationary or saddle point into a local minimum.

If a one-to-one linear mapping exists between the two parameter spaces $\{w_i\}$ and $\{c_i\}$, then a stationary point for the equalizer coefficients w_i must correspond to a stationary point in the c_i parameters. Consequently, the convergence properties in the c_i parameter space will be equivalent to those in the w_i parameter space. However, $c_i = \sum_{k=0}^{m} h_k w_{i-k}$ does not provide a one-to-one mapping. The linear mapping is one-to-one if and only if $c_i = \sum_{k=-\infty}^{\infty} h_k w_{i-k}$, i.e., the equalizer coefficients w_i must exist for $-\infty < i < \infty$.

In this case, the equalizer parameter vector $\mathbf{W}$ needs to be doubly infinite. Hence, *unless the equalizer has an infinite number of parameters and is infinitely non-causal*, the convergence behavior of the c_i parameters do not completely characterize the behavior of the finite-length equalizer [17].

Undesirable local convergence of the Godard algorithm to a high ISI equalizer was initially thought to be impossible due to some over-zealous interpretations of the global convergence results in the combined c_i space [15]. The local convergence of the Godard ($q = 2$) algorithm or CMA is accurately analyzed by Ding et al. [18], where it is shown that even for noiseless channels whose ISI can be completely eliminated by an FIR equalizer, there can be local convergence of this equalizer to undesirable minima of the cost surface. Furthermore, these equilibria still remain under moderate channel noise. Based on the convergence similarity between the Godard algorithm and the Shalvi-Weinstein algorithm, the local convergence of the Shalvi-Weinstein algorithm to undesirable minima is established in [14]. Using a similar method, Sato and BGR algorithms have also been seen to have additional local minima previously undetected in the combined parameter space [16].

The proof that existing blind equalization algorithms previously thought to be robust can converge to poor solutions demonstrates that rigorous convergence analyses of blind equalizers must be based on the actual equalizer coefficients. Moreover, the undesirable local convergence behavior of existing algorithms indicates the importance of algorithm parameter initialization, which can avoid these local convergent points.

24.7.3 Initialization Issues

In [19], it is shown that local minima of a CMA equalizer cost surface tend to exist near MMSE parameter settings if the delay δ is chosen to be too short or too long. In other words, convergence to local minima is more likely to occur when the equalizer has large tap weights concentrated near either end of the finite equalizer coefficient vector. This type of lopsided parameter weight distribution was also suggested in [15] as being indicative of a local convergence phenomenon. To avoid local convergence to a lopsided tap weight vector, Foschini [15] introduced a tap-centering initialization strategy that requires the gravity center of the equalizer coefficient vector be centered through periodic tap-shifting. A more recent result [14] shows that, by over-parameterization and tap-centering, the Godard algorithm or CMA can effectively reduce the probability of local convergence. This tap-centering method has also been proposed for the Shalvi-Weinstein algorithm [20].

In practice, the tap-centering initialization approach has become an integral part of most blind equalization algorithms. Although a thorough analysis of its effect has not been shown, most reported successful uses of blind equalization algorithms typically rely on tap-centering or center-spike initialization scheme [17]. Although special channels exist that can foil the successful convergence of Sato and BGR algorithms using tap-centering, such channels are atypical [16]. Hence, unless global convergence of the equalizer can be proven, tap-centering is commonly recommended for most blind equalizers.

24.8 Globally Convergent Equalizers

24.8.1 Linearly Constrained Equalizer With Convex Cost

Without a proof of global convergence and a thorough analysis on initialization of existing equalization methods, one can design new and possibly better blind algorithms that can proven to always result in the global minimization of ISI. Here we present one strategy based on highly specialized convex cost functions coupled with a constrained equalizer parameterization designed to avoid ill-convergence.

Recall that the goal of blind equalization is to remove ISI so that the equalizer output is

$$y(n) = ga(n-\delta), \quad g \neq 0 \,. \tag{24.37}$$

Blind equalization of pulse amplitude modulation (PAM) systems without gain recovery has been proposed in [21]. The idea is to fix the center tap w_0 as a non-zero constant in order to prevent equalizer to the trivial minimum with all zero coefficient values in a convex cost function. For QAM input, a nontrivial extension is shown here.

For the particular equalizer design, assume that the input QAM constellation is square, which resembles the constellation in Fig. 24.4(b). The cost function to be minimized is

$$J(\mathbf{W}) \stackrel{\Delta}{=} \max|\mathrm{Re}\{y(n)\}| = \max|Im\{y(n)\}| \,. \tag{24.38}$$

The convexity of $J(\mathbf{W})$ with respect to the equalizer coefficient vector $\mathbf{W}$ follows from the triangle inequality under the assumption that all input sequences are possible. We constrain the equalizer coefficients $\mathbf{W}(n)$ with the following linear constraint

$$\mathrm{Re}\{w_0\} + Im\{w_0\} = 1 \,, \tag{24.39}$$

where w_0 is the center tap. Due to the linearity of this constraint, the convexity of the cost function (24.38) with respect to both the real and imaginary parts of the equalizer coefficients is maintained, and global convergence is therefore assured.

Because of its convexity, this cost function is unimodal with a unique global minimum for almost all channels. It can then be shown [22] that a doubly infinite noncausal equalizer under the linear constraint is globally convergent to the condition in (24.37).

The linear constraint in (24.39) can be changed to any weighted linear combination of the two terms in (24.39). More general linear constraints on the equalizer coefficients can also be employed [23]. This fact is particularly important for preserving the global convergence property when causal finite-length equalizers are used. This behavior is a direct consequence of convexity, since restricting most of the equalizer taps to zero values as in FIR is a form of linear constraint. Convexity also ensures that one can approximate arbitrarily closely the performance of the ideal nonimplementable double infinite noncausal equalizer with a finite length FIR equalizer. These facts are important since many of the limitations illustrated earlier for convergence analyses of other equalizers can be overcome in this case.

For an actual implementation of this algorithm, a gradient descent method can be derived by using an l_p-norm cost function to approximate (24.38) as

$$J(\mathbf{W}) \approx E|\mathrm{Re}\{z_k\}|^p \, , \tag{24.40}$$

where p is a large integer. As the cost function in (24.40) is strictly convex, linear constraints such as truncation preserve convexity. Simulation examples of this algorithm can be found in [22] and [24].

24.9 Fractionally Spaced Blind Equalizers

A so-called fractionally spaced equalizer (FSE) is obtained from the system in Fig. 24.2 if the channel output is sampled at a rate faster than the baud or symbol rate $1/T$. Recent work on the blind FSE has been motivated by several new results on nonadaptive blind equalization based on second order cyclostationary statistics. In addition to the first noted work by Tong et al. [25], new nonadaptive algorithms are also presented in [26, 27, 28, 29]. Here we only focus on the adaptive framework.

Let p be an integer such that the sampling interval be $\Delta = T/p$. As long as the channel bandwidth is greater than the minimum $1/(2T)$, sampling at higher than $1/T$ can retain channel diversity as shown here.

Let the sequence of sampled channel output be

$$x(k\Delta) = \sum_{n=0}^{\infty} a(n)h(k\Delta - np\Delta + t_0) + w(k\Delta) \, . \tag{24.41}$$

For notational simplicity, the oversampled channel output $x(k\Delta)$ can be divided into p linearly-independent subsequences

$$x^{(i)}(n) \triangleq x[(np+i)\Delta] = x(nT + i\Delta), \qquad i = 1, \ldots, p \, . \tag{24.42}$$

Define K as the effective channel length based on

$$\begin{aligned} h_0^{(i)} &\neq 0, \quad \text{for some } 1 \le i \le p \\ h_K^{(i)} &\neq 0, \quad \text{for some } 1 \le i \le p \, . \end{aligned} \tag{24.43}$$

By denoting the sub-channel transfer function as

$$H_i(z) = \sum_{k=0}^{K} h_k^{(i)} z^{-k} \quad \text{where} \quad h_k^{(i)} \triangleq h(kT + i\Delta + t_0) \, , \tag{24.44}$$

the p subsequences can be written as

$$x^{(i)}(n) = H_i(z)a(n) + w(nT + i\Delta), \qquad i = 1, \ldots, p \, . \tag{24.45}$$

Thus, these p subsequences can be viewed as stationary outputs of p discrete FIR channels with a common input sequence $a(n)$ as shown in Fig. 24.5. Naturally, they can also represent physical sub-channels in multisensor receivers.

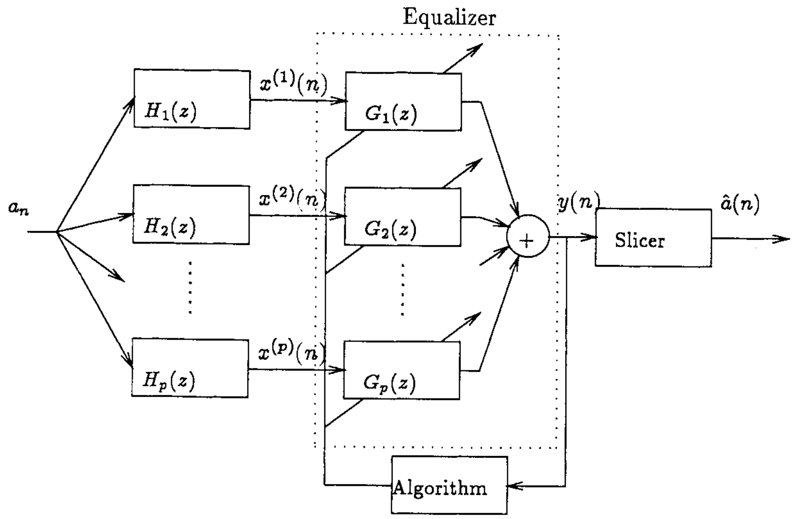

FIGURE 24.5: The structure of blind adaptive FSE.

The vector representation of the FSE is shown in Fig. 24.5. One equalizer filter is provided for each subsequence $x^{(i)}(n)$. In fact, the actual equalizer is a vector of filters

$$G_i(z) = \sum_{k=0}^{m} w_{i,k} z^{-k}, \qquad i = 1, \dots p\,. \tag{24.46}$$

The p filter outputs $\{y(n)^{(i)}\}$ are summed to form the stationary equalizer output

$$y(n) = \mathbf{W}^T \mathbf{X}(n) \tag{24.47}$$

where

$$\mathbf{W} \triangleq \left[w_{1,0} \;\; \dots \;\; w_{1,m} \;\; \dots w_{p,0} \;\; \dots \;\; w_{p,m}\right]^T .$$
$$\mathbf{X}(n) \triangleq \left[x(n)^{(1)} \;\; \dots \;\; x(n-m)^{(1)} \;\; \dots \;\; x(n)^{(p)} \;\; \dots \;\; x(n-m)^{(p)}\right]^T .$$

Given the equalizer output and parameter vector, any T-sampled blind equalization adaptive algorithm can be applied to the FSE via stochastic gradient descent techniques.

Since their first use, adaptive blind equalizers have often been implemented as FSEs. When training data are available, FSEs have the known advantage of suppressing timing phase sensitivity [30]. In fact, a blind FSE has another important advantage: there exists a one-to-one mapping between the combined parameter space and the equalizer parameter space, as shown in [31], under the following *length and zero conditions:*

- *The equalizer length satisfies* $(m+1) \geq K$*;*
- *The* p *discrete sub-channels* $\{H_i(z)\}$ *do not share any common zeros.*

Note that for T-sampled equalizers (TSE), only one ($p = 1$) sub-channel exists and all zeros are common zeros, and, thus, the length and zero conditions cannot be satisfied. In most practical implementations, p is either 2 or 3. So long as the above conditions hold, the convergence behaviors of blind adaptive FSEs can be characterized completely in the combined parameter space. Based on the work of [12, 15], for QAM channel inputs, there do not exist any algorithm-dependent stable equilibria other than the desired global minima (24.36) for FSEs driven by the Godard ($q = 2$) algorithm (CMA) and the Shalvi-Weinstein algorithms. Thus, the Godard and the Shalvi-Weinstein algorithms are globally convergent for FSEs satisfying these conditions [31].

Notice that global convergence of the Godard FSE is only proven for noiseless channels under the no-common zero condition. There have been recent advances in analyzing the performance of blind equalizers in the presence of Gaussian noise and the existence of common sub-channel zeros. While all possible delays of (24.36) are global minima for noiseless channels, the locations and effects of minima vary when channel noises are present. An analysis by Zeng and Tong shows that for noisy channels, CMA equalizer parameters have minima near the MMSE equilibria [32]. The effects of noise and common zeros was also studied by Fijalkow et al. [33, 34], providing further indications of the robustness of CMA when implemented as an FSE.

24.10 Concluding Remarks

Adaptive channel equalization and blind equalization are among the most successful applications of adaptive filtering. We have introduced the basic concept of blind equalization along with some of the most commonly used blind equalization algorithms. Without the aid of training signals, the key challenge of blind adaptive equalizers lies in the design of special cost functions whose minimization is consistent with the goal of ISI removal. We have also summarized key results on the convergence of blind equalizers The idea of constrained minimization of a convex cost function to assure global convergence of the blind equalizer was described. Finally, the blind adaptation in fractionally spaced equalizers and multichannel receivers was shown to possess useful convergence properties.

It is important to note that the problem of blind equalization has not been completely solved by any means. In addition to the fact that the convergence behaviors of most algorithms are still unknown, the rates of convergence of typical algorithms such as CMA is quite slow, often needing thousands of iterations to achieve acceptable output. The difficulty of the convergence analysis and the slow rate of convergence of these algorithms have prompted many efforts to modify blind error functions to obtain faster and better algorithms. Furthermore, nonadaptive algorithms that explicitly exploit higher order statistics [35]–[38] and second order cyclostationary statistics [25]–[29] appear to be quite efficient in exploiting small amount of channel output data. A detailed discussion of these methods is beyond the scope of the section. Interested readers may refer to two collected works edited by Haykin [24] and Gardner [39] and the references therein.

References

[1] Lucky, R.W., Techniques for adaptive equalization of digital communication systems, *Bell Syst. Tech. J.*, 45:255–286, Feb., 1966.

[2] Qureshi, S.U.H., Adaptive equalization, *Proc. IEEE*, 73:1349–1387, Sept., 1985.

[3] Benveniste, A., Goursat, M., and Ruget, G., Robust identification of a nonminimum phase system, *IEEE Trans. Auto. Control*, AC-25:385–399, June, 1980.

[4] Benveniste, A. and Goursat, M., Blind equalizers, *IEEE Trans. Commun.*, 32:871–882, Aug., 1982.

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

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

[7] Mazo, J.E., Analysis of decision-directed equalizer convergence, *Bell Syst. Tech. J.*, 59:1857–1876, Dec., 1980.

[8] Godard, D.N., Self-recovering equalization and carrier tracking in two-dimensional data communication systems, *IEEE Trans. Commun.*, COM-28:1867-1875, 1980.
[9] Treichler, J.R. and Agee, B.G., A new approach to multipath correction of constant modulus signals, *IEEE Trans. Acoustics, Speech Signal Process.*, ASSP-31:349-372, 1983.
[10] Picchi, G. and Prati, G., Blind equalization and carrier recovery using a "stop-and-go" decision-directed algorithm, *IEEE Trans. Commun.*, COM-35: 877-887, Sept., 1987.
[11] Hatzinakos, D., Blind equalization using stop-and-go criterion adaptation rules, *Opt. Eng.*, 31:1181-1198, June, 1992.
[12] Shalvi, O. and Weinstein, E., New criteria for blind deconvolution of nonminimum phase systems (channels), *IEEE Trans. Inform. Theory*, IT-36: 312-321, Mar., 1990.
[13] Brillinger, D.R. and Rosenblatt, M., Computation and interpretation of k-th order spectra, in *Spectral Analysis of Time Series*, B. Harris, Ed., Wiley, New York, 1967.
[14] Li, Y. and Ding, Z., Convergence analysis of finite length blind adaptive equalizers, *IEEE Trans. Signal Process.*, pp. 2120-2129, Sept., 1995.
[15] Foschini, G.J., Equalization without altering or detect data, *AT&T Tech. J.*, pp. 1885-1911, Oct., 1985.
[16] Ding, Z., Kennedy, R.A., Anderson, B.D.O., and Johnson, C.R., Jr., Local convergence of the Sato blind equalizer and generalizations under practical constraints, *IEEE Trans. Inform. Theory*, IT-39: 129-144, Jan., 1993.
[17] Ding, Z., Kennedy, R.A., and Johnson, C.R., Jr., On the (non)existence of undesirable equilibria of Godard blind equalizer, *IEEE Trans. Signal Process.*, 40:2425-2432, Oct., 1992.
[18] Ding, Z., Kennedy, R.A., Anderson, B.D.O., and Johnson, C.R., Jr., Ill-convergence of Godard blind equalizers in data communications, *IEEE Trans. Commun.*, 39:1313–1328, Sept., 1991.
[19] Minardi, M.J. and Ingram, M. Ann, Finding misconvergence in blind equalizers and new variance constraint cost functions to mitigate the problem, *Proc. 1996 Int. Conf. Acoustics, Speech, Signal Process.*, pp.1723-1726, May, 1996.
[20] Tugnait, J.K., Shalvi, O., and Weinstein, E., Comments on "new criteria for blind deconvolution of nonminimum phase systems (channels)," *IEEE Trans. Inf. Theory*, IT-38:210-213, Jan., 1992.
[21] Rupprecht, W.T., Adaptive equalization of binary NRZ-signals by means of peak value minimization, in *Proc. 7th Europ. Conf. Circuit Theory and Design*, pp. 352–355, Prague, 1985.
[22] Kennedy, R.A. and Ding, Z., Blind adaptive equalizers for QAM communication systems based on convex cost functions, *Opt. Eng.*, 31:1189-1199, June, 1992.
[23] Yamazaki, K. and Kennedy, R.A., Reformulation of linearly constrained adaptation and its application to blind equalization, *IEEE Trans. Signal Process.*, SP-42:1837-1841, 1994.
[24] Haykin, Simon, Ed., *Blind Deconvolution*, Prentice-Hall, 1994.
[25] Tong, L., Xu, G., and Kailath, T., Blind channel identification and equalization based on second-order statistics: a time-domain approach, *IEEE Trans. Inform. Theory*, IT-40:340-349, Mar., 1994.
[26] Moulines, E. et al., Subspace methods for the blind identification of multichannel FIR filters, *Proc. IEEE ICASSP*, pp. IV:573-576, Adelaide, 1994.
[27] Li, Y. and Ding, Z., ARMA system identification based on second order cyclostationarity, *IEEE Trans. Signal Process.*, pp. 3483-3493, Dec., 1994.
[28] Meriam, K.A., Duhamel, P., Gesbert, D., Loubaton, P., Mayrarague, S., Moulines, E., and Slock, D., Prediction error methods for time-domain blind identification of multichannel FIR filters, *Proc. IEEE Int. Conf. Acoustics, Speech, Signal Process.*, pp. 1968-1971, May, 1995.
[29] Hua, Y., Fast maximum likelihood for blind identification of multiple FIR channels, *IEEE Trans. Signal Process.*, SP-44:661-672, Mar., 1996.
[30] Gitlin, R.D. and Weinstein, S.B., Fractionally spaced equalization: an improved digital transversal equalizer, *Bell Syst. Tech. J.*, 60:275-296, 1981.

[31] Li, Y. and Ding, Z., Global convergence of fractionally spaced Godard adaptive equalizers, *IEEE Trans. Signal Process.*, SP-44:818-826, Apr., 1996.

[32] Zeng, H. and Tong, L., On the performance of CMA in the presence of noise, *Proc. Conf. Inform. Sci. Syst.*, Mar., 1996.

[33] Fijalkow, I., Treichler, J.R., and Johnson, C.R., Jr., Fractionally spaced blind equalization: loss of channel diversity, *Proc. IEEE Int. Conf. Acoustics, Speech, Signal Process.*, pp. 1988-1991, May, 1995.

[34] Touzni, A., Fijalkow, I., and Treichler, J.R., Fractionally spaced CMA under channel noise, *Proc. IEEE Int. Conf. Acoustics, Speech, and Signal Process.*, pp. 2674-2677, May, 1996.

[35] Tugnait, J.K., Identification of linear stochastic systems via second and fourth-order cumulant matching, *IEEE Trans. Inf. Theory*, IT-33:393-407, May, 1987.

[36] Giannakis, G.B. and Mendel, J.M., Identification of nonminimum phase systems using via higher order statistics, *IEEE Trans. Acoustics, Speech, and Signal Process.*, ASSP-37:360-377, 1989.

[37] Hatzinakos, D. and Nikias, C.L., Blind equalization using a tricepstrum based algorithm, *IEEE Trans. Commun.*, 39:669–682, May, 1991.

[38] Shalvi, O. and Weinstein, E., Super-exponential methods for blind deconvolution, *IEEE Trans. Inform. Theory*, IT-39:504-519, Mar., 1993.

[39] Gardner, W.A., Ed., *Cyclostationarity in Communications and Signal Processing*, IEEE Press, 1994.