

# Adaptive Blind Channel Equalization

Shafayat Abrar<sup>1</sup>, Azzedine Zerguine<sup>2</sup> and Asoke Kumar Nandi<sup>3</sup>

<sup>1</sup>*Department of Electrical Engineering,  
COMSATS Institute of Information Technology, Islamabad 44000*

<sup>2</sup>*Department of Electrical Engineering,  
King Fahd University of Petroleum and Minerals, Dhahran 31261*

<sup>3</sup>*Department of Electrical Engineering and Electronics,  
The University of Liverpool, Liverpool L69 3BX*

<sup>1</sup>*Pakistan*

<sup>2</sup>*Saudi Arabia*

<sup>3</sup>*United Kingdom*

## 1. Introduction

For bandwidth-efficient communication systems, operating in high inter-symbol interference (ISI) environments, adaptive equalizers have become a necessary component of the receiver architecture. An accurate estimate of the amplitude and phase distortion introduced by the channel is essential to achieve high data rates with low error probabilities. An adaptive equalizer provides a simple practical device capable of both learning and inverting the distorting effects of the channel. In conventional equalizers, the filter tap weights are initially set using a training sequence of data symbols known both to the transmitter and receiver. These trained equalizers are effective and widely used. Conventional least mean square (LMS) adaptive filters are usually employed in such supervised receivers, see Haykin (1996).

However, there are several drawbacks to the use of training sequences. Implementing a training sequence can involve significant transceiver complexity. Like in a point-to-multipoint network transmissions, sending training sequences is either impractical or very costly in terms of data throughput. Also for slowly varying channels, an initial training phase may be tolerable. However, there are scenarios where training may not be feasible, for example, in equalizer implementations of digital cellular handsets. When the communications environment is highly non-stationary, it may even become grossly impractical to use training sequences. A blind equalizer, on the other hand, does not require a training sequence to be sent for start-up or restart. Rather, the blind equalization algorithms use a priori knowledge regarding the statistics of the transmitted data sequence as opposed to an exact set of symbols known both to the transmitter and receiver. In addition, the specifications of training sequences are often left ambiguous in standards bodies, leading to vendor specific training sequences and inter-operability problems. Blind equalization solves this problem as well, see Ding & Li (2001); Garth et al. (1998); Haykin (1994).

In this Chapter, we provide an introduction to the basics of adaptive blind equalization. We describe popular methodologies and criteria for designing adaptive algorithms for blind

equalization. Most importantly, we discuss how to use the probability density function (PDF) of transmitted signal to design ISI-sensitive cost functions. We discuss the issues of admissibility of proposed cost function and stability of derived adaptive algorithm.

## 2. Trained and blind adaptive equalizers: Historical perspectives

Adaptive trained channel equalization was first developed by Lucky for telephone channels, see Lucky (1965; 1966). Lucky proposed the so-called zero-forcing (ZF) method to be applied in FIR equalization. It was an adaptive procedure and in a noiseless situation, the optimal ZF equalizer tends to be the inverse of the channel. In the mean time, Widrow and Hoff introduced the least mean square (LMS) adaptive algorithm which begins adaptation with the aid of a training sequence known to both transmitter and receiver, see Widrow & Hoff (1960); Widrow et al. (1975). The LMS algorithm is capable of reducing mean square error (MSE) between the equalizer output and the training sequence. Once the signal eye is open, the equalizer is then switched to tracking mode which is commonly known as decision-directed mode. The decision-directed method is unsupervised and its effectiveness depends on the initial condition of equalizer coefficients; if the initial eye is closed then it is likely to diverge.

In blind equalization, the desired signal is unknown to the receiver, except for its probabilistic or statistical properties over some known alphabets. As both the channel and its input are unknown, the objective of blind equalization is to recover the unknown input sequence based solely on its probabilistic and statistical properties, see C.R. Johnson, Jr. et al. (1998); Ding & Li (2001); Haykin (1994). Historically, the possibility of blind equalization was first discussed in Allen & Mazo (1974), where the authors proved analytically that an adjusting equalizer, optimizing the mean-squared sample values at its output while keeping a particular tap anchored at unit value, is capable of inverting the channel without needing a training sequence. In the subsequent year, Sato was the first who came up with a robust realization of an adaptive blind equalizer for PAM signals, see Sato (1975). It was followed by a number of successful attempts on blind magnitude equalization (i.e., equalization without carrier-phase recovery) in Godard (1980) for complex-valued signals (V29/QPSK/QAM), in Treichler & Agee (1983) for AM/FM signals, in Serra & Esteves (1984) and Bellini (1986) for PAM signals. However, many of these algorithms originated from intuitive starting points.

The earliest works on joint blind equalization and carrier-phase recovery were reported in Benveniste & Goursat (1984); Kennedy & Ding (1992); Picchi & Prati (1987); Wesolowski (1987). Recent references include Abrar & Nandi (2010a;b;c;d); Abrar & Shah (2006a); Abrar & Qureshi (2006b); Abrar et al. (2005); Goupil & Palicot (2007); Im et al. (2001); Yang et al. (2002); Yuan & Lin (2010); Yuan & Tsai (2005). All of these blind equalizers are capable of recovering the true power of transmitted data upon convergence and are classified as *Bussgang*-type, see Bellini (1986). The Bussgang blind equalization algorithms make use of a nonlinear estimate of the channel input. The memoryless nonlinearity, which is the function of equalizer output, is designed to minimize an ISI-sensitive cost function that implicitly exploits higher-order statistics. The performance of such kind of blind equalizer depends strongly on the choice of nonlinearity.

The first comprehensive analytical study of the blind equalization problem was presented by Benveniste, Goursat, and Ruget in Benveniste et al. (1980a;b). They established that if

the transmitted signal is composed of non-Gaussian, independent and identically distributed samples, both channel and equalizer are linear time-invariant filters, noise is negligible, and the probability density functions of transmitted and equalized signals are equal, then the channel has been perfectly equalized. This mathematical result is very important since it establishes the possibility of obtaining an equalizer with the sole aid of signal's statistical properties and without requiring any knowledge of the channel impulse response or training data sequence. Note that the very term "blind equalization" can be attributed to Benveniste and Goursat from the title of their paper Benveniste & Goursat (1984). This seminal paper 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 into a class of algorithms based on non-MSE cost functions. More importantly, the convergence properties of the proposed algorithms were carefully investigated.

The second analytical landmark occurred in 1990 when Shalvi and Weinstein significantly simplified the conditions for blind equalization, see Shalvi & Weinstein (1990). Before this work, it was usually believed that one needs to exploit infinite statistics to ensure zero-forcing equalization. Shalvi and Weinstein showed that the zero-forcing equalization can be achieved if only two statistics of the involved signals are restored. Actually, they proved that, if the fourth order cumulant (kurtosis) is maximized and the second order cumulant (energy) remains the same, then the equalized signal would be a scaled and rotated version of the transmitted signal. Interesting accounts on Shalvi and Weinstein criterion can be found in Tugnait et al. (1992) and Romano et al. (2011).

### 3. System model and "Bussgang" blind equalizer

The baseband model for a typical complex-valued data communication system consists of an unknown linear time-invariant channel  $\{h\}$  which represents the physical inter-connection between the transmitter and the receiver. The transmitter generates a sequence of complex-valued random input data  $\{a_n\}$ , each element of which belongs to a complex alphabet  $\mathcal{A}$ . The data sequence  $\{a_n\}$  is sent through the channel whose output  $x_n$  is observed by the receiver. The input/output relationship of the channel can be written as:

$$x_n = \sum_k a_{n-k} h_k + v_n, \quad (1)$$

where the additive noise  $v_n$  is assumed to be stationary, Gaussian, and independent of the channel input  $a_n$ . We also assume that the channel is stationary, moving-average and has finite length. The function of equalizer at the receiver is to estimate the delayed version of original data,  $a_{n-\delta}$ , from the received signal  $x_n$ . Let  $\mathbf{w}_n = [w_{n,0}, w_{n,1}, \dots, w_{n,N-1}]^T$  be vector of equalizer coefficients with  $N$  elements (superscript  $T$  denotes transpose). Let  $\mathbf{x}_n = [x_n, x_{n-1}, \dots, x_{n-N+1}]^T$  be the vector of channel observations. The output of the equalizer is

$$y_n = \mathbf{w}_n^H \mathbf{x}_n \quad (2)$$

where superscript  $H$  denotes conjugate transpose. If  $\{t\} = \{h\} \star \{w^*\}$  represents the overall channel-equalizer impulse response (where  $\star$  denotes convolution), then (2) can be expressed

as:

$$y_n = \sum_i w_i^* x_{n-i} = \sum_l a_{n-l} t_l + v_n' = t_\delta a_{n-\delta} + \underbrace{\sum_{l \neq \delta} t_l a_{n-l}}_{\text{signal} + \text{ISI} + \text{noise}} + v_n' \quad (3)$$

Equation (3) distinctly exposes the effects of multi-path inter-symbol interference and additive noise. Even in the absence of additive noise, the second term can be significant enough to cause an erroneous detection.

The idea behind a Bussgang blind equalizer is to minimize (or maximize), through the choice of equalizer filter coefficients  $w$ , a certain cost function  $J$ , depending on the equalizer output  $y_n$ , such that  $y_n$  provides an estimate of the source signal  $a_n$  up to some inherent indeterminacies, giving,  $y_n = \alpha a_{n-\delta}$ , where  $\alpha = |\alpha|e^{j\gamma} \in \mathbb{C}$  represents an arbitrary gain. The phase  $\gamma$  represents an isomorphic rotation of the symbol constellation and hence is dependent on the rotational symmetry of signal alphabets; for example,  $\gamma = m\pi/2$  radians, with  $m \in \{0, 1, 2, 3\}$  for a quadrature amplitude modulation (QAM) system. Hence, a Bussgang blind equalizer tries to solve the following optimization problem:

$$w^\dagger = \arg \text{optimize}_w J, \text{ with } J = E[\mathcal{J}(y_n)] \quad (4)$$

The cost  $J$  is an expression for implicitly embedded higher-order statistics of  $y_n$  and  $\mathcal{J}(y_n)$  is a real-valued function. Ideally, the cost  $J$  makes use of statistics which are significantly modified as the signal propagates through the channel, and the optimization of cost modifies the statistics of the signal at the equalizer output, aligning them with those at channel input. The equalization is accomplished when equalized sequence  $y_n$  acquires an identical distribution as that of the channel input  $a_n$ , see Benveniste et al. (1980a). If the implementation method is realized by stochastic gradient-based adaptive approach, then the updating algorithm is

$$w_{n+1} = w_n \pm \mu \left( \frac{\partial \mathcal{J}}{\partial w_n} \right)^* \quad (5a)$$

$$= w_n + \mu \Phi(y_n)^* x_n, \text{ with } \Phi(y_n) = \pm \frac{\partial \mathcal{J}}{\partial y_n^*} \quad (5b)$$

where  $\mu$  is step-size, governing the speed of convergence and the level of steady-state performance, see Haykin (1996). The positive and negative signs in (5a) are respectively for maximization and minimization. The complex-valued error-function  $\Phi(y_n)$  can be understood as an estimate of the difference between the desired and the actual equalizer outputs. That is,  $\Phi(y_n) = \Psi(y_n) - y_n$ , where  $\Psi(y_n)$  is an estimate of the transmitted data  $a_n$ . The nonlinear memory-less estimate,  $\Psi(y_n)$ , is usually referred to as Bussgang nonlinearity and is selected such that, at steady state, when  $y_n$  is close to  $a_{n-\delta}$ , the autocorrelation of  $y_n$  becomes equal to the cross-correlation between  $y_n$  and  $\Psi(y_n)$ , i.e.,

$$E[y_n \Phi(y_{n-i})^*] = 0 \Rightarrow E[y_n y_{n-i}^*] = E[y_n \Psi(y_{n-i})^*]$$

An admissible estimate of  $\Psi(y_n)$ , however, is the conditional expectation  $E[a_{n'}|y_n]$ , see Nikias & Petropulu (1993). Using Bayesian estimation technique,  $E[a_{n'}|y_n]$  was derived in

Bellini (1986); Fiori (2001); Haykin (1996); Pinchas & Bobrovsky (2006; 2007). These methods, however, rely on explicit computation of higher-order statistics and are not discussed here.

#### 4. Trained and blind equalization design methodologies

Generally, a blind equalization algorithm attempts to invert the channel using both the received data samples and certain known statistical properties of the input data. For example, it is easy to show that for a minimum phase channel, the spectra of the input and output signals of the channel can be used to determine the channel impulse response. However, most communication channels do not possess minimum phase. To identify a non-minimum phase channel, a non-Gaussian signal is required along with nonlinear processing at the receiver using higher-order moments of the signal, see Benveniste et al. (1980a;b). Based upon available analysis, simulations, and experiments in the literature, it can be said that an admissible blind cost function has two main attributes: **1)** it makes use of statistics which are significantly modified as the signal propagates through the channel, and **2)** optimization of the cost function modifies the statistics of the signal at the channel output, aligning them with the statistics of the signal at the channel input.

Designing a blind equalization cost function has been lying strangely more in the realm of art than science; majority of the cost functions tend to be proposed on intuitive grounds and then validated. Due to this reason, a plethora of blind cost functions is available in literature. On the contrary, the fact is that there exist established methods which facilitate the designing of blind cost functions requiring statistical properties of transmitted and received signals. One of the earliest methods originated in late 70's in geophysics community who sought to determine the inverse of the channel in seismic data analysis and it was named minimum entropy deconvolution (MED), see Gray (1979b); Wiggins (1977). Later in early 90's, Satorius and Mulligan employed MED principle and came up with several proposals to blindly equalize the communication channels, see Satorius & Mulligan (1993). However, those marvelous signal-specific proposals regrettably failed to receive serious attention.

In the sequel, we discuss MED along with other popular methods for designing blind cost functions and corresponding adaptive equalizers.

##### 4.1 Lucky criterion

In 1965, Lucky suggested that the propagation channel may be inverted by an equalizer if equalizer minimizes the following peak distortion criterion, see Lucky (1965):

$$J_{\text{peak}} = \frac{1}{|t_\delta|} \sum_{l \neq \delta} |t_l| \quad (6)$$

This criterion is equivalent to requiring that the equalizer maximizes the eye opening. the intuitive explanation of (6) is as follows. From (3), ignoring the noise, obtain the value of error  $\mathcal{E}$  due to ISI, given as  $\mathcal{E} = y_n - a_{n-\delta} = a_{n-\delta} (t_\delta - 1) + \sum_{l \neq \delta} t_l a_{n-l}$ . Assuming the maximum and minimum values of  $a_n$  are  $\tilde{a}$  and  $-\tilde{a}$ , respectively; the maximum error is easily written as

$$\mathcal{E}_{\text{max}} = |y_n - a_{n-\delta}|_{\text{max}} = |a_{n-\delta}| |t_\delta - 1| + \tilde{a} \sum_{l \neq \delta} |t_l| \quad (7)$$

If  $t_\delta$  is close to unity, then the cost  $J_{\text{peak}}$  is the scaled version of  $\mathcal{E}_{\text{max}}$  as given by  $J_{\text{peak}} \approx \mathcal{E}_{\text{max}}/(|t_\delta|\tilde{a})$ . It is important to note that the cost  $J_{\text{peak}}$  is a convex function of the equalizer weights and has a well-defined minimum. Thus any minimum of  $J_{\text{peak}}$  found by gradient search or other systematic programming methods must be the absolute (or global) minimum of distortion, see Lucky (1966). To prove the convexity of  $J_{\text{peak}}$ , it is necessary to show that for two equalizer settings  $\mathbf{w}^{(a)}$  and  $\mathbf{w}^{(b)}$  and for all  $\eta, 0 \leq \eta \leq 1$

$$J_{\text{peak}}(\eta \mathbf{w}^{(a)} + (1 - \eta) \mathbf{w}^{(b)}) \leq \eta J_{\text{peak}}(\mathbf{w}^{(a)}) + (1 - \eta) J_{\text{peak}}(\mathbf{w}^{(b)}) \quad (8)$$

The above equation shows that the distortion always lies on or beneath the chord joining values of distortion in  $N$ -spaces. Below is the proof of (8):

$$\begin{aligned} J_{\text{peak}}(\eta \mathbf{w}^{(a)} + (1 - \eta) \mathbf{w}^{(b)}) &= \sum_{l \neq \delta} \left| \sum_k h_k (\eta (w_{l-k}^{(a)})^* + (1 - \eta) (w_{l-k}^{(b)})^*) \right| \\ &\leq \eta \sum_{l \neq \delta} \left| \sum_k h_k (w_{l-k}^{(a)})^* \right| + (1 - \eta) \sum_{l \neq \delta} \left| \sum_k h_k (w_{l-k}^{(b)})^* \right| \\ &= \eta J_{\text{peak}}(\mathbf{w}^{(a)}) + (1 - \eta) J_{\text{peak}}(\mathbf{w}^{(b)}). \end{aligned} \quad (9)$$

However, in practice, it is not an easy one to achieve this convexity in a gradient search (adaptive) procedure, as it is necessary to obtain the projection of the gradient onto the constraint hyperplane, see Allen & Mazo (1973). Alternatively one can also seek to minimize the mean-square distortion criterion:

$$J_{\text{ms}} = \frac{1}{|t_\delta|^2} \sum_{l \neq \delta} |t_l|^2 \quad (10)$$

The cost  $J_{\text{ms}}$  is not convex but unimodal, mathematically tractable and capable of yielding admissible solution, see Allen & Mazo (1973). Under the assumption  $|t_\delta| = \max\{|t_l|\}$ , Shalvi & Weinstein (1990) used the expression (10) to quantify ISI, i.e.,  $\text{ISI} = J_{\text{ms}}$ . Using criterion (10), we can formulate the following tractable problem for ISI mitigation:

$$\mathbf{w}^\dagger = \arg \min_{\mathbf{w}} \sum_{l \neq \delta} |t_l|^2 \quad \text{s.t.} \quad |t_\delta| = 1. \quad (11)$$

where we have assumed that the equalizer coefficients ( $\mathbf{w}^\dagger$ ) have been selected such that the condition ( $|t_\delta| = 1$ ) is always satisfied. Introducing the channel autocorrelation matrix  $\mathbf{H}$ , whose  $(i, j)$  element is given by  $\mathcal{H}_{ij} = \sum_k h_{k-i} h_{k-j}^*$ , we can show that  $\sum_l |t_l|^2 = \mathbf{w}_n^H \mathbf{H} \mathbf{w}_n$ . The equalizer has to make one of the coefficients of  $\{t_l\}$  say  $t_\delta = t_0 = \mathbf{w}_n^H \mathbf{h}$  to be unity and others to be zero, where  $\mathbf{h} = [h_{K-1}, h_{K-2}, \dots, h_1, h_0]^T$ ; it gives the value of ISI of an unequalized system at time index  $n$  as follows:

$$\text{ISI} = \frac{\mathbf{w}_n^H \mathbf{H} \mathbf{w}_n}{|\mathbf{w}_n^H \mathbf{h}|^2} - 1. \quad (12)$$

Now consider the optimization of the problem (11). Using Lagrangian multiplier  $\lambda$ , we obtain

$$\sum_{l \neq \delta} |t_l|^2 + \lambda (t_\delta - 1) = \mathbf{w}_n^H \mathbf{H} \mathbf{w}_n - 1 + \lambda (\mathbf{w}_n^H \mathbf{h} - 1) \quad (13)$$

Next differentiating with respect to  $\mathbf{w}_n^*$  and equating to zero, we get  $\mathbf{H} \mathbf{w}_n + \lambda \mathbf{h} = 0 \Rightarrow \mathbf{w}_n = -\lambda \mathbf{H}^{-1} \mathbf{h}$ . Substituting the above value of  $\mathbf{w}_n$  in (12), we obtain

$$\text{ISI}_n = \frac{\mathbf{h}^H \mathbf{H}^{-1} \mathbf{h}}{|\mathbf{h}^H \mathbf{H}^{-1} \mathbf{h}|^2} - 1. \quad (14)$$

To appreciate the possible benefit of solution (14), consider a channel  $h_{-1} = 1 - \varepsilon$ ,  $h_0 = \varepsilon$  and  $h_1 = 0$ , where  $0 \leq \varepsilon \leq 1$ . Without equalizer, we have

$$\text{ISI} = \frac{(1 - \varepsilon)^2 + \varepsilon^2}{\max(1 - \varepsilon, \varepsilon)^2} - 1. \quad (15)$$

The ISI approaches zero when  $\varepsilon$  is either zero or unity. Assuming a 2-tap equalizer, we obtain

$$\mathbf{H} = \begin{bmatrix} (1 - \varepsilon)^2 + \varepsilon^2 & (1 - \varepsilon)\varepsilon \\ (1 - \varepsilon)\varepsilon & (1 - \varepsilon)^2 + \varepsilon^2 \end{bmatrix} \quad (16)$$

Using (14) and (16), we obtain

$$\text{ISI} = \frac{\varepsilon^2(1 - \varepsilon)^2}{1 - 4\varepsilon + 6\varepsilon^2 - 4\varepsilon^3 + 2\varepsilon^4}. \quad (17)$$

Refer to Fig. 1, the ISI (17) of the equalized system is lower than that of the uncompensated system. The adaptive implementation of  $J_{\text{ms}}$  can be realized in a supervised scenario.

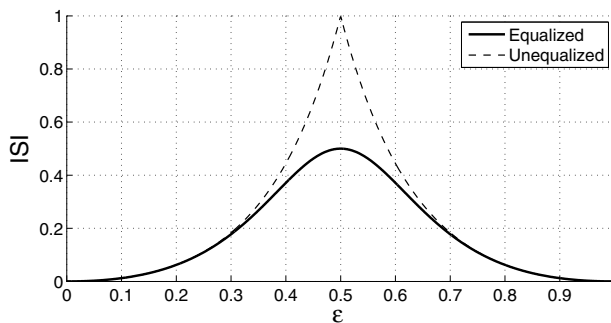


Fig. 1. ISI of unequalized and equalized systems.

Combining the two expression (7) and (10), the following cost is obtained which is usually termed as mean square error (MSE) criterion, see Widrow et al. (1975):

$$J_{\text{mse}} = \text{E} \left[ |a_{n-\delta} - y_n|^2 \right] \quad (18)$$

Minimizing (18), we obtain the following update:

$$\mathbf{w}_{n+1} = \mathbf{w}_n + \mu (a_{n-\delta} - y_n)^* \mathbf{x}_n, \quad (19)$$

which is known as LMS algorithm or Widrow-Hoff algorithm. Note that the parameter  $\mu$  is a positive step-size and the following value of  $\mu \equiv \mu_{\text{LMS}}$  ensures the stability of algorithm, see Farhang-Boroujeny (1998):

$$0 < \mu_{\text{LMS}} < \frac{1}{2NP_a}, \quad (20)$$

where  $P_a = E[|a|^2]$  is the average energy of signal  $a_n$  and  $N$  is the length of equalizer.

## 4.2 Minimum entropy deconvolution criterion

The minimum entropy deconvolution (MED) is probably the earliest principle for designing cost functions for blind equalization. This principle was introduced by Wiggins in seismic data analysis in the year 1977, who sought to determine the inverse channel  $\mathbf{w}^\dagger$  that maximizes the *kurtosis* of the deconvolved data  $y_n$ , see Wiggins (1977; 1978). For seismic data, which are super-Gaussian in nature, he suggested to maximize the following cost:

$$\frac{\frac{1}{B} \sum_{b=1}^B |y_{n-b+1}|^4}{\left( \frac{1}{B} \sum_{b=1}^B |y_{n-b+1}|^2 \right)^2} \quad (21)$$

This deconvolution scheme seeks the smallest number of large spikes consistent with the data, thus maximizing the order or, equivalently, minimizing the entropy or disorder in the data, Walden (1985). Note that the equation (21) has the statistical form of sample kurtosis and the expression is scale-invariant. Later, in the year 1979, Gray generalized the Wiggins' proposal with two degrees of freedom as follows, Gray (1979b):

$$J_{\text{med}}^{(p,q)} \equiv \frac{\frac{1}{B} \sum_{b=1}^B |y_{n-b+1}|^p}{\left( \frac{1}{B} \sum_{b=1}^B |y_{n-b+1}|^q \right)^{\frac{p}{q}}} \quad (22)$$

The criterion was rigorously investigated in Donoho (1980), where Donoho developed general rules for designing MED-type estimators. Several cases of MED, in the context of blind deconvolution of seismic data, have appeared in the literature, like  $J_{\text{med}}^{(2,1)}$  in Ooe & Ulrych (1979),  $J_{\text{med}}^{(4,2)}$  in Wiggins (1977),  $\lim_{\epsilon \rightarrow 0} J_{\text{med}}^{(p-\epsilon, p)}$  in Claerbout (1977),  $J_{\text{med}}^{(p,2)}$  in Gray (1978), and  $J_{\text{med}}^{(2p,p)}$  in Gray (1979a).

In the derivation of the criterion (22), it is assumed that the original signal  $a_n$ , which is primary reflection coefficients in geophysical system or transmitted data in communication systems, can be modeled as realization of independent non-Gaussian process with distribution

$$p_{\mathcal{A}}(a; \alpha) = \frac{\alpha}{2\beta\Gamma\left(\frac{1}{\alpha}\right)} \exp\left(-\frac{|a|^\alpha}{\beta^\alpha}\right) \quad (23)$$

where signal  $a_n$  is real-valued,  $\alpha$  is the shape parameter,  $\beta$  is the scale parameter, and  $\Gamma(\cdot)$  is the Gamma function. This family covers a wide range of distributions. The certain event



( $\alpha = 0$ ), double exponential ( $\alpha = 1$ ), Gaussian ( $\alpha = 2$ ), and uniform distributions ( $\alpha \rightarrow \infty$ ) are all members. For geophysical deconvolution problem, we have the range  $0.6 \leq \alpha \leq 1.5$ , and for communication system where the signals are uniformly distributed we have ( $\alpha \rightarrow \infty$ ). Although, signals in communication are discrete, the equation (23) is still good to approximate some densely and uniformly distributed signal.

In the context of geophysics, where the primary coefficient  $a_n$  is super-Gaussian, maximizing the criterion (23) drives the distribution of deconvolved sequence  $y_n$  away from  $p_Y(y_n; p)$  towards  $p_Y(y_n; q)$ , where  $p > q$ . However, for the communication blind equalization problem, the underlying distribution of the transmitted (possibly pulse amplitude modulated) data symbols are closer to a uniform density (sub-Gaussian) and thus we would minimize the cost (23) with  $p > q$ . We have the following cost for blind equalization of communication channel:

$$w^+ = \begin{cases} \arg \min_w J_{\text{med}}^{(p,q)}, & \text{if } p > q, \\ \arg \max_w J_{\text{med}}^{(p,q)}, & \text{if } p < q. \end{cases} \quad (24)$$

The feasibility of (24) for blind equalization of digital signals has been studied in Satorius & Mulligan (1992; 1993) and Benedetto et al. (2008). In Satorius & Mulligan (1992), implementing (24) with  $p > q$ , the following adaptive algorithm was obtained:

$$w_{n+1} = w_n + \mu \sum_{k=1}^{\hat{B}} \left( \frac{\sum_{b=1}^B |y_{n-b+1}|^p}{\sum_{b=1}^B |y_{n-b+1}|^q} |y_{n-k+1}|^{q-2} - |y_{n-k+1}|^{p-2} \right) y_{n-k+1}^* x_{n-k+1}, \quad (25)$$

In the sequel, we will refer to (25) as Satorius-Mulligan algorithm (SMA). Also, for a detailed discussion on the stochastic approximate realization of MED, refer to Walden (1988).

### 4.3 Constant modulus criterion

The most popular and widely studied blind equalization criterion is the constant modulus criterion, Godard (1980); Treichler & Agee (1983); Treichler & Larimore (1985); it is given by

$$J_{\text{cm}} = E \left[ (|y_n|^2 - R_{\text{cm}})^2 \right], \quad (26)$$

where  $R_{\text{cm}} = E[|a|^4] / E[|a|^2]$  is a statistical constant usually termed as dispersion constant. For an input signal that has a constant modulus  $|a_n| = \sqrt{R_{\text{cm}}}$ , the criterion penalizes output samples  $y_n$  that do not have the desired constant modulus characteristics. This modulus restoral concept has a particular advantage in that it allows the equalizer to be adapted independent of carrier recovery. Because the cost is insensitive to the phase of  $y_n$ , the equalizer adaptation can occur independently and simultaneously with the operation of the carrier recovery system. This property also makes it applicable to analog modulation signals with constant amplitude such as those using frequency or phase modulation, see Treichler & Larimore (1985). The stochastic gradient-descent minimization of (26) yields the following algorithm:

$$w_{n+1} = w_n + \mu (R_{\text{cm}} - |y_n|^2) y_n^* x_n, \quad (27)$$

which is usually termed as constant modulus algorithm (CMA). Note that, considering  $\hat{B} = 1$ ,  $p = 4$ ,  $q = 2$  and large  $B$ , SMA (25) reduces to CMA.

If data symbols are independent and identically distributed, noise is negligible and the length of the equalizer is infinite then after some calculations, the CM cost may be expressed as a function of joint channel-equalizer impulse response coefficients as follows:

$$J_{\text{cm}} = \left( \mathbb{E}[|a_n|^4] - 2\mathbb{E}[|a_n|^2]^2 \right) \sum_l |t_l|^4 + 2\mathbb{E}[|a_n|^2]^2 \left( \sum_l |t_l|^2 \right)^2 - 2\mathbb{E}[|a_n|^4] \sum_l |t_l|^2 + \text{const.} \quad (28)$$

As in Godard (1980), the partial derivative of  $J_{\text{cm}}$  with respect to  $t_k$  can be written as

$$\frac{\partial J_{\text{cm}}}{\partial t_k} = 4t_k \left( \mathbb{E}[|a_n|^4] (|t_k|^2 - 1) + 2\mathbb{E}[|a_n|^2]^2 \sum_{l \neq k} |t_l|^2 \right) \quad (29)$$

The minimum can be found by solving  $\frac{\partial J_{\text{cm}}}{\partial t_k} = 0$ , i.e.,

$$t_k \left( \mathbb{E}[|a_n|^4] (|t_k|^2 - 1) + 2\mathbb{E}[|a_n|^2]^2 \sum_{l \neq k} |t_l|^2 \right) = 0, \quad \forall k \quad (30)$$

Unfortunately, the set of equations has an infinite number of solutions; the cost  $J_{\text{cm}}$  is thus non-convex. The solutions  $T_{\mathcal{M}}$ ,  $\mathcal{M} = 1, 2, \dots$ , can be represented as follows: all elements of the set  $\{t_l\}$  are equal to zero, except  $\mathcal{M}$  of them and those non-zero elements have equal magnitude of  $\sigma_{\mathcal{M}}^2$  defined by

$$\sigma_{\mathcal{M}}^2 = \frac{\mathbb{E}[|a_n|^4]}{\mathbb{E}[|a_n|^4] + 2(\mathcal{M} - 1)\mathbb{E}[|a_n|^2]^2} \quad (31)$$

Among these solutions, under the condition  $\mathbb{E}[|a_n|^4] < 2\mathbb{E}[|a_n|^2]^2$ , the solution  $T_1$  is that for which the energy is the largest at the equalizer output and ISI is zero. The absolute minimum of  $J_{\text{cm}}$  is therefore reached in the case of zero ISI.

#### 4.4 Shtrom-Fan criterion

In the year 1998, Shtrom and Fan presented a class of cost functions for achieving blind equalization which were solely the function of  $\{t\}$  parameters, see Shtrom & Fan (1998). They suggested to minimize the difference between any two norms of the joint channel-equalizer impulse response, each raised to the same power, i.e.,

$$J_{\text{sf}} = \left( \sum_l |t_l|^p \right)^{r/p} - \left( \sum_l |t_l|^q \right)^{r/q}, \quad p < q \text{ and } r > 0 \quad (32)$$

where  $p, q, r \in \mathbb{R}$ . This proposal was based on the following property of vector norms:

$$\lim_{s \rightarrow 0} \sqrt[s]{\sum_l |t_l|^s} \geq \dots \geq \sqrt[p]{\sum_l |t_l|^p} \geq \sqrt[q]{\sum_l |t_l|^q} \geq \dots \geq \lim_{m \rightarrow \infty} \sqrt[m]{\sum_l |t_l|^m} \quad (33)$$

where  $p < q$  and equality occurs if and only if  $t_l = \pm \delta_{l-k}$ ,  $k \in \mathbb{Z}$ , which is precisely the zero-forcing condition. From the above there is a multitude of cost functions to choose from. From (32), we have following possibilities to minimize:

$$\sum_l |t_l| - \max_l \{|t_l|\}, \quad (p = 1, q \rightarrow \infty, r = 1) \quad (34a)$$

$$(\sum_l |t_l|)^2 - \sum_l |t_l|^2, \quad (p = 1, q = r = 2) \quad (34b)$$

$$(\sum_l |t_l|^2)^2 - \sum_l |t_l|^4, \quad (p = 2, q = r = 4) \quad (34c)$$

$$(\sum_l |t_l|^2)^3 - \sum_l |t_l|^6, \quad (p = 2, q = r = 6) \quad (34d)$$

$$(\sum_l |t_l|^4)^2 - \sum_l |t_l|^8, \quad (p = 4, q = r = 8). \quad (34e)$$

Some of these cost functions are easily implementable, whereas others are not.

Consider  $p = 2$  and  $q = r = 2m$  in (32) to obtain a subclass:

$$J_{sf}^{sub} = \left( \sum_l |t_l|^2 \right)^m - \sum_l |t_l|^{2m} \quad (35)$$

This subclass is not convex, although it is potentially unimodal in  $\mathbf{t}$  domain and easily implementable. As in Shtrom & Fan (1998), the partial derivative of  $J_{sf}^{sub}$  with respect to  $t_k$  can be written as

$$\frac{\partial J_{sf}^{sub}}{\partial t_k} = 2mt_k \left( \left( \sum_l |t_l|^2 \right)^{m-1} - |t_k|^{2(m-1)} \right) \quad (36)$$

The equation (36) has two solutions, one of which corresponds to  $t_l = 0$ ,  $\forall l$ . This solution will not occur if a constraint is imposed. The other solution is the minimum corresponding to zero-forcing condition. This is seen from (36) as  $\sum_l |t_l|^2 = |t_k|^2$ , which can only hold when  $\mathbf{t}$  has at most one nonzero element, i.e., the desired delta function. Now compare this result with that of constant modulus in equation (31) which contains multiple nonzero-forcing solutions. It means, in contrast to CMA, it is less likely to have local minima in  $J_{sf}^{sub}$  in equalizer domain.

The cost functions (34a), in their current form, are not directly applicable in real scenario as we have no information of  $\{t\}$ 's. These costs need to be converted from functions of  $\{t\}$ 's to functions of  $y_n$ 's. As in Shtrom & Fan (1998), we can show that

$$\sum_l |t_l|^2 = \frac{C_{1,1}^{y_n}}{C_{1,1}^a} = \frac{E[|y_n|^2]}{E[|a|^2]} \quad (37a)$$

$$\sum_l |t_l|^4 = \frac{C_{2,2}^{y_n}}{C_{2,2}^a} = \frac{E[|y_n|^4] - 2E[|y_n|^2]^2}{E[|a|^4] - 2E[|a|^2]^2} \quad (37b)$$

$$\sum_l |t_l|^6 = \frac{C_{3,3}^{y_n}}{C_{3,3}^a} = \frac{E[|y_n|^6] - 9E[|y_n|^4]E[|y_n|^2] + 12E[|y_n|^2]^3}{E[|a|^6] - 9E[|a|^4]E[|a|^2] + 12E[|a|^2]^3} \quad (37c)$$

where  $C_{p,q}^z$  is  $(p+q)$ th order cumulant of complex random variable defined as follows:

$$C_{p,q}^z = \text{cumulant} \left( \underbrace{z_1, \dots, z_p}_{p \text{ terms}}; \underbrace{z_1^*, \dots, z_q^*}_{q \text{ terms}} \right) \quad (38)$$

Using (37a) and assuming  $m = 2$ , we obtain the following expression for  $J_{sf}^{\text{sub}}$ :

$$J_{sf}^{\text{sub}} = \left( \frac{E[|y_n|^2]}{E[|a|^2]} \right)^2 - \frac{E[|y_n|^4] - 2E[|y_n|^2]^2}{E[|a|^4] - 2E[|a|^2]^2} \quad (39)$$

Minimizing (39) with respect to coefficients  $w^*$ , we obtain the following adaptive algorithm:

$$w_{n+1} = w_n + \mu \left( \frac{R_{cm}}{P_a} \hat{E}[|y_n|^2] - |y_n|^2 \right) y_n^* x_n, \quad (40a)$$

$$\hat{E}[|y_{n+1}|^2] = \hat{E}[|y_n|^2] + \frac{1}{n} \left( |y_n|^2 - \hat{E}[|y_n|^2] \right) \quad (40b)$$

where  $P_a$  is the average energy of signal  $a_n$  and  $R_{cm}$  is the same statistical constant as we defined in CMA. Note that the algorithm requires an iterative estimate of equalizer output energy. We will refer to (40) as Shtrom-Fan Algorithm (SFA).

#### 4.5 Shalvi-Weinstein criterion

Earlier to Shtrom and Fan, in the year 1990, Shalvi and Weinstein suggested a criterion that laid the theoretical foundation to the problem of blind equalization, see Shalvi & Weinstein (1990). They demonstrated that the condition of equality between the PDF's of the transmitted and equalized signals, due to BGR theorem Benveniste et al. (1980a,b), was excessively tight. Under the similar assumptions, as laid by Benveniste *et al.*, they demonstrated that it is possible to perform blind equalization by satisfying the condition  $E[|y_n|^2] = E[|a_n|^2]$  and ensuring that a nonzero cumulant of order higher than 2 of  $a_n$  and  $y_n$  are equal.

For a two dimensional signal  $a_n$  with four-quadrant symmetry (i.e.,  $E[a_n^2] = 0$ ), they suggested to maximize the following unconstrained cost function (which involved second and fourth order cumulants):

$$J_{sw} = \text{sgn}[C_{2,2}^{a_n}] \left( C_{2,2}^{y_n} + (\gamma_1 + 2) (C_{1,1}^{y_n})^2 + 2\gamma_2 C_{1,1}^{y_n} \right) \quad (41a)$$

$$= \text{sgn}[C_{2,2}^{a_n}] \left( E[|y_n|^4] + \gamma_1 E[|y_n|^2]^2 + 2\gamma_2 E[|y_n|^2] \right) \quad (41b)$$

where  $\gamma_1$  and  $\gamma_2$  are some statistical constants. The corresponding stochastic gradient algorithm is given by

$$w_{n+1} = w_n + \mu \text{sgn}[C_{2,2}^{a_n}] \left( \gamma_1 \hat{E}[|y_n|^2] + |y_n|^2 + \gamma_2 \right) y_n^* x_n, \quad (42a)$$

$$\hat{E}[|y_{n+1}|^2] = \hat{E}[|y_n|^2] + \frac{1}{n} \left( |y_n|^2 - \hat{E}[|y_n|^2] \right) \quad (42b)$$

where ( $\text{sgn}[C_{2,2}^{a_n}] = -1$ ) due to the sub-Gaussian nature of digital signals. The above algorithm is usually termed as Shalvi-Weinstein algorithm (SWA). Note that SWA unifies CMA and SFA (i.e., the specific case in Equation (40)). Substituting  $\gamma_1 = 0$  and  $\gamma_2 = -R_{cm}$  in SWA, we obtain CMA (27). Similarly, substituting  $\gamma_2 = 0$  and  $\gamma_1 = -R_{cm}/P_a$  in SWA, we obtain SFA (40). Note that the Shtrom-Fan criterion appears to be the generalization of Shalvi-Weinstein criterion with cumulants of generic orders.

## 5. Blind equalization of APSK signal: Employing MED principle

### 5.1 Designing a blind cost function

We employ MED principle and use the PDFs of transmitted amplitude-phase shift-keying (APSK) and ISI-affected received signal to design a cost function for blind equalization. Consider a continuous APSK signal, where signal alphabets  $\{a_R + j a_I\} \in \mathcal{A}$  are assumed to be uniformly distributed over a circular region of radius  $R_a$  and center at the origin. The joint PDF of  $a_R$  and  $a_I$  is given by (refer to Fig. 3(a))

$$p_{\mathcal{A}}(a_R + j a_I) = \begin{cases} \frac{1}{\pi R_a^2}, & \sqrt{a_R^2 + a_I^2} \leq R_a, \\ 0, & \text{otherwise.} \end{cases} \quad (43)$$

Now consider the transformation  $\mathcal{Y} = \sqrt{a_R^2 + a_I^2}$  and  $\Theta = \angle(a_R, a_I)$ , where  $\mathcal{Y}$  is the modulus and  $\angle()$  denotes the angle in the range  $(0, 2\pi)$  that is defined by the point  $(i, j)$ . The joint distribution of the modulus  $\mathcal{Y}$  and  $\Theta$  can be obtained as  $p_{\mathcal{Y}, \Theta}(\tilde{y}, \tilde{\theta}) = \tilde{y} / (\pi R_a^2)$ ,  $\tilde{y} \geq 0$ ,  $0 \leq \tilde{\theta} < 2\pi$ . Since  $\mathcal{Y}$  and  $\Theta$  are independent, we obtain a triangular distribution for  $\mathcal{Y}$  given by  $p_{\mathcal{Y}}(\tilde{y} : H_0) = 2\tilde{y} / R_a^2$ ,  $\tilde{y} \geq 0$ , where  $H_0$  denotes the hypothesis that signal is distortion-free. Let  $\mathcal{Y}_n, \mathcal{Y}_{n-1}, \dots, \mathcal{Y}_{n-N+1}$  be a sequence, of size  $N$ , obtained by taking modulus of randomly

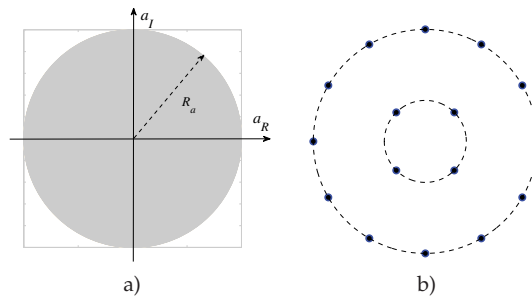


Fig. 2. a) A continuous APSK, and b) a discrete practical 16APSK.

generated distortion-free signal alphabets  $\mathcal{A}$ , where subscript  $n$  indicates discrete time index. Let  $\mathcal{Z}_1, \mathcal{Z}_2, \dots, \mathcal{Z}_N$  be the order statistic of sequence  $\{\mathcal{Y}\}$ . Let  $p_{\mathcal{Y}}(\tilde{y}_n, \tilde{y}_{n-1}, \dots, \tilde{y}_{n-N+1} : H_0)$  be an  $N$ -variate density of the continuous type, then, under the hypothesis  $H_0$ , we obtain

$$p_{\mathcal{Y}}(\tilde{y}_n, \tilde{y}_{n-1}, \dots, \tilde{y}_{n-N+1} : H_0) = \frac{2^N}{R_a^{2N}} \prod_{k=1}^N \tilde{y}_{n-k+1}. \quad (44)$$

Next we find  $p_{\mathcal{Y}}^*(\tilde{y}_n, \tilde{y}_{n-1}, \dots, \tilde{y}_{n-N+1} : H_0)$  as follows:

$$\begin{aligned} p_{\mathcal{Y}}^*(\tilde{y}_n, \tilde{y}_{n-1}, \dots, \tilde{y}_{n-N+1} : H_0) &= \int_0^\infty p_{\mathcal{Y}}(\lambda \tilde{y}_n, \lambda \tilde{y}_{n-1}, \dots, \lambda \tilde{y}_{n-N+1} : H_0) \lambda^{N-1} d\lambda \\ &= \frac{2^N}{R_a^{2N}} \prod_{k=1}^N \tilde{y}_{n-k+1} \int_0^{R_a/\tilde{z}_N} \lambda^{2N-1} d\lambda = \frac{2^{N-1}}{N (\tilde{z}_N)^{2N}} \prod_{k=1}^N \tilde{y}_{n-k+1}, \end{aligned} \quad (45)$$

where  $\tilde{z}_1, \tilde{z}_2, \dots, \tilde{z}_N$  are the order statistic of elements  $\tilde{y}_n, \tilde{y}_{n-1}, \dots, \tilde{y}_{n-N+1}$ , so that  $z_1 = \min\{\tilde{y}\}$  and  $z_N = \max\{\tilde{y}\}$ . Now consider the next hypothesis ( $H_1$ ) that signal suffers with multi-path interference as well as with additive Gaussian noise (refer to Fig. 3(b)). Due to which, the in-phase and quadrature components of the received signal are modeled as normal distributed; owing to central limit theorem, it is theoretically justified. It means that the modulus of the received signal follows Rayleigh distribution,

$$p_Y(\tilde{y} : H_1) = \frac{\tilde{y}}{\sigma_{\tilde{y}}^2} \exp\left(-\frac{\tilde{y}^2}{2\sigma_{\tilde{y}}^2}\right), \tilde{y} \geq 0, \sigma_{\tilde{y}} > 0. \quad (46)$$

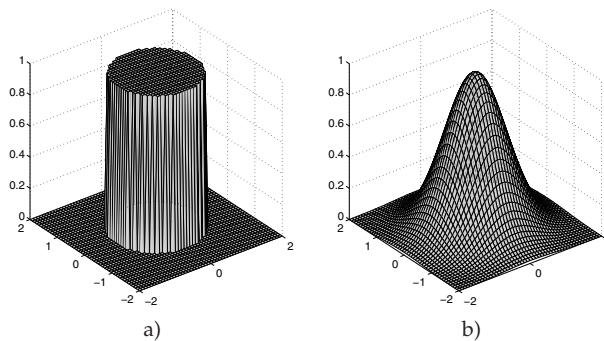


Fig. 3. PDFs (not to scale) of a) continuous APSK and b) Gaussian distributed received signal.

The  $N$ -variate densities  $p_Y(\tilde{y}_n, \tilde{y}_{n-1}, \dots, \tilde{y}_{n-N+1} : H_1)$  and  $p_Y^*(\tilde{y}_n, \tilde{y}_{n-1}, \dots, \tilde{y}_{n-N+1} : H_1)$  are obtained as

$$p_Y(\tilde{y}_n, \tilde{y}_{n-1}, \dots, \tilde{y}_{n-N+1} : H_1) = \frac{1}{\sigma_{\tilde{y}}^{2N}} \prod_{k=1}^N \tilde{y}_{n-k+1} \exp\left(-\frac{\tilde{y}_{n-k+1}^2}{2\sigma_{\tilde{y}}^2}\right) \quad (47)$$

$$p_Y^*(\tilde{y}_n, \tilde{y}_{n-1}, \dots, \tilde{y}_{n-N+1} : H_1) = \frac{\prod_{k=1}^N \tilde{y}_{n-k+1}}{\sigma_{\tilde{y}}^{2N}} \int_0^\infty \exp\left(-\frac{\lambda^2 \sum_{k=1}^N \tilde{y}_{n-k+1}^2}{2\sigma_{\tilde{y}}^2}\right) \lambda^{2N-1} d\lambda \quad (48)$$

Substituting  $u = \frac{1}{2} \lambda^2 \sigma_{\tilde{y}}^{-2} \sum_{k=1}^N \tilde{y}_{n-k+1}^2$ , we obtain

$$p_Y^*(\tilde{y}_n, \tilde{y}_{n-1}, \dots, \tilde{y}_{n-N+1} : H_1) = \frac{2^{N-1} \Gamma(N)}{\left(\sum_{k=1}^N \tilde{y}_{n-k+1}^2\right)^N} \prod_{k=1}^N \tilde{y}_{n-k+1} \quad (49)$$

The scale-invariant uniformly most powerful test of  $p_Y^*(\tilde{y}_n, \tilde{y}_{n-1}, \dots, \tilde{y}_{n-N+1} : H_0)$  against  $p_Y^*(\tilde{y}_n, \tilde{y}_{n-1}, \dots, \tilde{y}_{n-N+1} : H_1)$  provides us, see Sidak et al. (1999):

$$O(\tilde{y}_n) = \frac{p_Y^*(\tilde{y}_n, \tilde{y}_{n-1}, \dots, \tilde{y}_{n-N+1} : H_0)}{p_Y^*(\tilde{y}_n, \tilde{y}_{n-1}, \dots, \tilde{y}_{n-N+1} : H_1)} = \frac{1}{N!} \left[ \frac{\sum_{k=1}^N \tilde{y}_{n-k+1}^2}{\tilde{z}_N^2} \right]^N \underset{H_1}{\overset{H_0}{\geq}} C \quad (50)$$

where  $C$  is a threshold. Assuming large  $N$ , we can approximate  $\frac{1}{N} \sum_{k=1}^N \tilde{y}_{n-k+1}^2 \approx \mathbb{E}[|y_n|^2]$ . It helps obtaining a statistical cost for the blind equalization of APSK signal as follows:

$$w^\dagger = \arg \max_w \frac{\mathbb{E}[|y_n|^2]}{(\max\{|y_n|\})^2} \quad (51)$$

Based on the previous discussion, maximizing cost (51) can be interpreted as determining the equalizer coefficients,  $w$ , which drives the distribution of its output,  $y_n$ , away from Gaussian distribution toward uniform, thus removing successfully the interference from the received APSK signal. Note that the above result (51) may be obtained directly from (24) by substituting  $p = 2$  and  $q \rightarrow \infty$ , see Abrar & Nandi (2010b).

## 5.2 Admissibility of the proposed cost

The cost (51) demands maximizing equalizer output energy while minimizing the largest modulus. Since the largest modulus of transmitted signal  $a_n$  is  $R_a$ , incorporating this *a priori* knowledge, the unconstrained cost (51) can be written in a constrained form as follows:

$$w^\dagger = \arg \max_w \mathbb{E}[|y_n|^2] \quad \text{s.t.} \quad \max\{|y_n|\} \leq R_a. \quad (52)$$

By incorporating  $R_a$ , it would be possible to recover the true energy of signal  $a_n$  upon successful convergence. Also note that  $\max\{|y_n|\} = R_a \sum_l |t_l|$  and  $\mathbb{E}[|y_n|^2] = P_a \sum_l |t_l|^2$ . Based on which, we note that the cost (52) is quadratic, and the feasible region (constraint) is a convex set (proof of which is similar to that in Equation (9)). The problem, however, is non-convex and may have multiple local maxima. Nevertheless, we have the following theorem:

**Theorem:** Assume  $w^\dagger$  is a local optimum in (52), and  $t^\dagger$  is the corresponding total channel equalizer impulse-response and channel noise is negligible. Then it holds  $|t_l| = \delta_{l-l^\dagger}$ .

**Proof:** Without loss of generality we assume that the channel and equalizer are real-valued. We re-write (52) as follows:

$$w^\dagger = \arg \max_w \sum_l t_l^2 \quad \text{s.t.} \quad \sum_l |t_l| \leq 1. \quad (53)$$

Now consider the following quadratic problem in  $t$  domain

$$t^\dagger = \arg \max_t \sum_l t_l^2 \quad \text{s.t.} \quad \sum_l |t_l| \leq 1. \quad (54)$$

Assume  $t^{(f)}$  is a feasible solution to (54). We have

$$\sum_l t_l^2 \leq \left( \sum_l |t_l| \right)^2 \leq 1 \quad (55)$$

and

$$\left( \sum_l |t_l| \right)^2 = \sum_l t_l^2 + \sum_{l_1} \sum_{l_2, l_2 \neq l_1} |t_{l_1} t_{l_2}| \quad (56)$$

The first equality in (55) is achieved if and only if all cross terms in (56) are zeros. Now assume that  $\mathbf{t}^{(k)}$  is a local optimum of (54), i.e., the following proposition holds

$$\exists \varepsilon > 0, \forall \mathbf{t}^{(f)}, \|\mathbf{t}^{(f)} - \mathbf{t}^{(k)}\|_2 \leq \varepsilon \quad (57)$$

$\Rightarrow \sum_l (t_l^{(k)})^2 \geq \sum_l (t_l^{(f)})^2$ . Suppose  $\mathbf{t}^{(k)}$  does not satisfy the **Theorem**. Consider  $\mathbf{t}^{(c)}$  defined by

$$t_{l_1}^{(c)} = t_{l_1}^{(k)} + \frac{\varepsilon}{\sqrt{2}},$$

$$t_{l_2}^{(c)} = t_{l_2}^{(k)} - \frac{\varepsilon}{\sqrt{2}},$$

and  $t_l^{(c)} = t_l^{(k)}$ ,  $l \neq l_1, l_2$ . We also assume that  $t_{l_2}^{(k)} < t_{l_1}^{(k)}$ . Next, we have  $\|\mathbf{t}^{(c)} - \mathbf{t}^{(k)}\|_2 = \varepsilon$ , and  $\sum_l |t_l^{(c)}| = \sum_l |t_l^{(k)}| \leq 1$ . However, one can observe that

$$\sum_l (t_l^{(k)})^2 - \sum_l (t_l^{(c)})^2 = \sqrt{2}\varepsilon (t_{l_2}^{(k)} - t_{l_1}^{(k)}) - \varepsilon^2 < 0, \quad (58)$$

which means  $\mathbf{t}^{(k)}$  is not a local optimum to (54). Therefore, we have shown by a counterexample that all local maxima of (54) should satisfy the **Theorem**.

### 5.3 Adaptive optimization of the proposed cost

For a stochastic gradient-based adaptive implementation of (52), we need to modify it to involve a *differentiable* constraint; one of the possibilities is

$$\mathbf{w}^\dagger = \arg \max_{\mathbf{w}} \mathbb{E} [|y_n|^2] \quad \text{s.t.} \quad \text{fmax}(R_a, |y_n|) = R_a, \quad (59)$$

where we have used the following identity (below  $a, b \in \mathbb{C}$ ):

$$\text{fmax}(|a|, |b|) \equiv \frac{|a| + |b| + ||a| - |b||}{2} = \begin{cases} |a|, & \text{if } |a| \geq |b| \\ |b|, & \text{otherwise.} \end{cases} \quad (60)$$

The function  $\text{fmax}$  is differentiable, viz

$$\frac{\partial \text{fmax}(|a|, |b|)}{\partial a^*} = \frac{a(1 + \text{sgn}(|a| - |b|))}{4|a|} = \begin{cases} a/(2|a|), & \text{if } |a| > |b| \\ 0, & \text{if } |a| < |b| \end{cases} \quad (61)$$

If  $|y_n| < R_a$ , then the cost (59) simply maximizes output energy. However, if  $|y_n| > R_a$ , then the constraint is violated and the new update  $\mathbf{w}_{n+1}$  is required to be computed such that the magnitude of a *a posteriori* output  $\mathbf{w}_{n+1}^H \mathbf{x}_n$  becomes smaller than or equal to  $R_a$ . Next, employing Lagrangian multiplier, we get

$$\mathbf{w}^\dagger = \arg \max_{\mathbf{w}} \left\{ \mathbb{E} [|y_n|^2] + \lambda (\text{fmax}(R_a, |y_n|) - R_a) \right\}. \quad (62)$$

The stochastic approximate gradient-based optimization of  $\mathbf{w}^\dagger = \arg \max_{\mathbf{w}} \mathbb{E} [\mathcal{J}]$  is realized as  $\mathbf{w}_{n+1} = \mathbf{w}_n + \mu \partial \mathcal{J} / \partial \mathbf{w}_n^*$ , where  $\mu > 0$  is a small step-size. Differentiating (62) with respect



to  $w_n^*$  gives

$$\frac{\partial |y_n|^2}{\partial w_n^*} = \frac{\partial |y_n|^2}{\partial y_n} \frac{\partial y_n}{\partial w_n^*} = y_n^* x_n$$

and

$$\frac{\partial \text{fmax}(R_a, |y_n|)}{\partial w_n^*} = \frac{\partial \text{fmax}(R_a, |y_n|)}{\partial y_n} \frac{\partial y_n}{\partial w_n^*} = \frac{g_n y_n^*}{4|y_n|} x_n,$$

where  $g_n \equiv 1 + \text{sgn}(|y_n| - R_a)$ ; we obtain  $w_{n+1} = w_n + \mu(1 + \lambda g_n / (4|y_n|)) y_n^* x_n$ . If  $|y_n| < R_a$ , then  $g_n = 0$  and  $w_{n+1} = w_n + \mu y_n^* x_n$ . Otherwise, if  $|y_n| > R_a$ , then  $g_n = 2$  and

$$w_{n+1} = w_n + \mu(1 + \lambda / (2|y_n|)) y_n^* x_n.$$

As mentioned earlier, in this case, we have to compute  $\lambda$  such that  $w_{n+1}^H x_n$  lies inside the circular region without sacrificing the output energy. Such an update can be realized by minimizing  $|y_n|^2$  while satisfying the Bussgang condition, see Bellini (1994). Note that the satisfaction of Bussgang condition ensures recovery of the true signal energy upon successful convergence. One of the possibilities is  $\lambda = -2(1 + \beta)|y_n|$ ,  $\beta > 0$ , which leads to

$$w_{n+1} = w_n + \mu(-\beta) y_n^* x_n.$$

The Bussgang condition requires

$$\underbrace{\text{E}[y_n y_{n-i}^*]}_{|y_n| < R_a} + \underbrace{(-\beta) \text{E}[y_n y_{n-i}^*]}_{|y_n| > R_a} = 0, \quad \forall i \in \mathbb{Z} \quad (63)$$

In steady-state, we assume  $y_n = a_{n-\delta} + u_n$ , where  $u_n$  is convolutional noise. For  $i \neq 0$ , (63) is satisfied due to uncorrelated  $a_n$  and independent and identically distributed samples of  $u_n$ . Let  $a_n$  comprises  $M$  distinct symbols on  $L$  moduli  $\{R_1, \dots, R_L\}$ , where  $R_L = R_a$  is the largest modulus. Let  $M_i$  denote the number of unique (distortion-free) symbols on the  $i$ th modulus, i.e.,  $\sum_{l=1}^L M_l = M$ . With negligible  $u_n$ , we solve (63) for  $i = 0$  to get

$$M_1 R_1^2 + \dots + M_{L-1} R_{L-1}^2 + \frac{1}{2} M_L R_L^2 - \frac{\beta}{2} M_L R_L^2 = 0 \quad (64)$$

The last two terms indicate that, when  $|y_n|$  is close to  $R_L$ , it would be equally likely to update in either direction. Noting that  $\sum_{l=1}^L M_l R_l^2 = M P_a$ , the simplification of (64) gives

$$\beta = 2 \frac{M}{M_L} \frac{P_a}{R_a^2} - 1. \quad (65)$$

The use of (65) ensures recovery of true signal energy upon successful convergence. Finally the proposed algorithm is expressed as

$$w_{n+1} = w_n + \mu f(y_n) y_n^* x_n, \quad (66)$$

$$f(y_n) = \begin{cases} 1, & \text{if } |y_n| \leq R_a \\ -\beta, & \text{if } |y_n| > R_a. \end{cases}$$

Note that the error-function  $f(y_n)y_n^*$  has 1) finite derivative at the origin, 2) is increasing for  $|y_n| < R_a$ , 3) decreasing for  $|y_n| > R_a$  and 4) insensitive to phase/frequency offset errors. In Baykal et al. (1999), these properties 1)-4) have been regarded as essential features of a constant modulus algorithm; this motivates us to denote (66) as  $\beta$ CMA.

#### 5.4 Stability of the derived adaptive algorithm

In this Section, we carry out a deterministic stability analysis of  $\beta$ CMA for any bounded magnitude received signal. The analysis relies on the analytical framework of Rupp & Sayed (2000). We shall assume that the successive regression vectors  $\{x_i\}$  are nonzero and also uniformly bounded from above and from below. We update the equalizer only when its output amplitude is higher than certain threshold; we stop the update otherwise.

In our analysis, we assume that the threshold is  $R_a$ . So we only consider those updates when  $|y_n| > R_a$ ; we extract and denote the active update steps with time index  $k$ . We study the following form:

$$w_{k+1} = w_k + \mu_k \Phi_k^* x_k, \quad \Phi_k \neq 0, \quad k = 0, 1, 2, \dots \quad (67)$$

where  $\Phi_k \equiv \Phi(y_k) = f(y_k)y_k$ . Let  $w_*$  denote vector of the optimal equalizer and let  $z_k = w_*^H x_k = a_{k-\delta}$  is the optimal output so that  $|z_k| = R_a$ . Define the *a priori* and *a posteriori* estimation errors

$$\begin{aligned} e_k^a &:= z_k - y_k = \bar{w}_k^H x_k \\ e_k^p &:= z_k - s_k = \bar{w}_{k+1}^H x_k \end{aligned} \quad (68)$$

where  $\bar{w}_k = w_* - w_k$ . We assume that  $|e_k^a|$  is small and equalizer is operating in the vicinity of  $w_*$ . We introduce a function  $\zeta(x, y)$ :

$$\zeta(x, y) := \frac{\Phi(y) - \Phi(x)}{x - y} = \frac{f(y)y - f(x)x}{x - y}, \quad (x \neq y) \quad (69)$$

Using  $\zeta(x, y)$  and simple algebra, we obtain

$$\Phi_k = f(y_k) y_k = \zeta(z_k, y_k) e_k^a \quad (70)$$

$$e_k^p = \left( 1 - \frac{\mu_k}{\bar{\mu}_k} \zeta(z_k, y_k) \right) e_k^a \quad (71)$$

where  $\bar{\mu}_k = 1/\|x_k\|^2$ . For the stability of adaptation, we require  $|e_k^p| < |e_k^a|$ . To ensure it, we require to guarantee the following for all possible combinations of  $z_k$  and  $y_k$ :

$$\left| 1 - \frac{\mu_k}{\bar{\mu}_k} \zeta(z_k, y_k) \right| < 1, \quad \forall k \quad (72)$$

Now we require to prove that the real part of the function  $\zeta(z_k, y_k)$  defined by (69) is positive and bounded from below. Recall that  $|z_k| = R_a$  and  $|y_k| > R_a$ . We start by writing  $z_k/y_k = re^{j\phi}$  for some  $r < 1$  and for some  $\phi \in [0, 2\pi)$ . Then expression (69) leads to

$$\zeta(z_k, y_k) = \frac{\overbrace{f(y_k)}^{(-\beta)} y_k - \overbrace{f(z_k)}^{(=0)} z_k}{z_k - y_k} = \frac{\beta y_k}{y_k - z_k} = \frac{\beta}{1 - re^{j\phi}}. \quad (73)$$

It is important for our purpose to verify whether the real part of  $\beta/(1 - re^{j\phi})$  is positive. For any fixed value of  $r$ , we allow the angle  $\phi$  to vary from zero to  $2\pi$ , then the term  $\beta/(1 - re^{j\phi})$  describes a circle in the complex plane whose least positive value is  $\beta/(1 + r)$ , obtained for  $\phi = \pi$ , and whose most positive value is  $\beta/(1 - r)$ , obtained for  $\phi = 0$ . This shows that for  $r \in (0, 1)$ , the real part of the function  $\xi(z_k, y_k)$  lies in the interval

$$\frac{\beta}{1+r} \leq \xi_R(z_k, y_k) \leq \frac{\beta}{1-r} \quad (74)$$

Referring to Fig. 4, note that the function  $\xi(z_k, y_k)$  assumes values that lie inside a circle in the

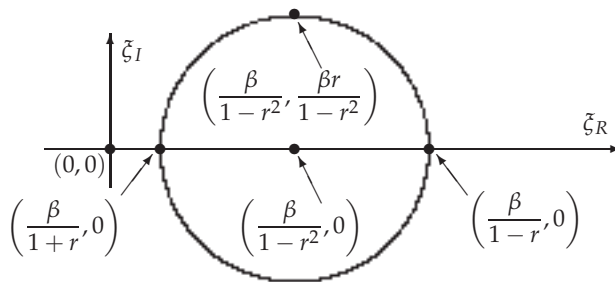


Fig. 4. Plot of  $\xi(z_k, y_k)$  for arbitrary  $\beta, r$  and  $\phi \in [0, 2\pi)$ .

right-half plane. From this figure, we can obtain the following bound for the imaginary part of  $\xi(z_k, y_k)$  (that is  $\xi_I(z_k, y_k)$ ):

$$-\frac{\beta r}{1-r^2} \leq \xi_I(z_k, y_k) \leq \frac{\beta r}{1-r^2}. \quad (75)$$

Let  $A$  and  $B$  be any two positive numbers satisfying

$$A^2 + B^2 < 1. \quad (76)$$

We need to find a  $\mu_k$  that satisfies

$$\left| \frac{\mu_k}{\bar{\mu}_k} \xi_I(z_k, y_k) \right| < A \Rightarrow \mu_k < \frac{A \bar{\mu}_k}{|\xi_I(z_k, y_k)|} \quad (77)$$

and

$$\left| 1 - \frac{\mu_k}{\bar{\mu}_k} \xi_R(z_k, y_k) \right| < B \Rightarrow \mu_k > \frac{(1-B)\bar{\mu}_k}{\xi_R(z_k, y_k)} \quad (78)$$

Combining (77) and (78), we obtain

$$0 < \frac{(1-B)\bar{\mu}_k}{\xi_R(z_k, y_k)} < \mu_k < \frac{A \bar{\mu}_k}{|\xi_I(z_k, y_k)|} < 1 \quad (79)$$

Using the extremum values of  $\xi_R(z_k, y_k)$  and  $\xi_I(z_k, y_k)$ , we obtain

$$\frac{(1+r)(1-B)}{\beta \|x_k\|^2} < \mu_k < \frac{(1-r^2)A}{\beta r \|x_k\|^2} \quad (80)$$

We need to guarantee that the upper bound in the above expression is larger than the lower bound. This can be achieved by choosing  $\{A, B\}$  properly such that

$$0 < (1 - B) < \frac{1 - r}{r} A < 1 \quad (81)$$

From our initial assumptions that the equalizer is in the vicinity of open-eye solution and  $|y_k| > R_a$ , we know that  $r < 1$ . It implies that we require to determine the *smallest* value of  $r$  which satisfies (81), or in other words, we have to determine the bound for step-size with the *largest* equalizer output amplitude. In Fig. 5(a), we plot the function  $f_r := (1 - r)/r$  versus  $r$ ; note that for  $0.5 \leq r < 1$  we have  $f_r \leq 1$ .

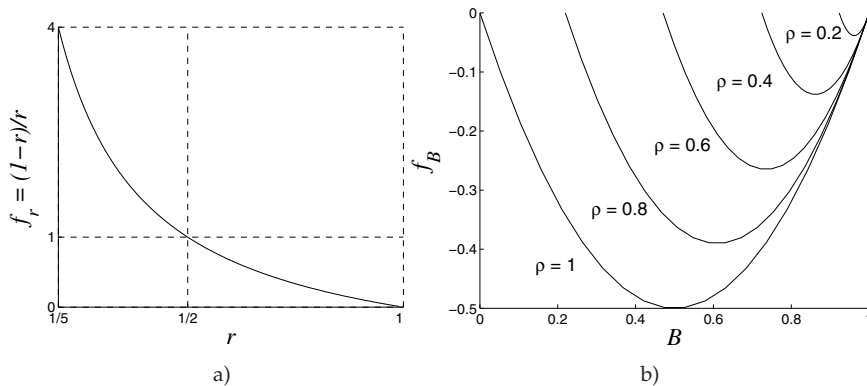


Fig. 5. Plots: a)  $f_r$  and b)  $f_B$ .

Let  $\{A_o, B_o\}$  be such that  $(1 - B_o) = \rho A_o$ , where  $0 < \rho < 1$ . To satisfy (81), we need  $0.5 \leq r < 1$  and  $\rho < (1 - r)/r$ . From (76),  $B_o$  must be such that

$$(1 - B_o)^2 \rho^{-2} + B_o^2 < 1 \quad (82)$$

which reduces to the following quadratic inequality in  $B_o$ :

$$(1 + \rho^{-2}) B_o^2 - 2\rho^{-2} B_o + (\rho^{-2} - 1) < 0. \quad (83)$$

If we find a  $B_o$  that satisfies this inequality, then a pair  $\{A_o, B_o\}$  satisfying (76) and (81) exists. So consider the quadratic function  $f_B := (1 + \rho^{-2}) B^2 - 2\rho^{-2} B + (\rho^{-2} - 1)$ . It has a negative minimum and it crosses the real axis at the positive roots  $B^{(1)} = (1 - \rho^2) / (1 + \rho^2)$ , and  $B^{(2)} = 1$ . This means that there exist many values of  $B$ , between the roots, at which the quadratic function in  $B$  evaluates to negative values (refer to Fig. 5(b)).

Hence,  $B_o$  falls in the interval  $(1 - \rho^2) / (1 + \rho^2) < B_o < 1$ ; it further gives  $A_o = 2\rho / (1 + \rho^2)$ . Using  $\{A_o, B_o\}$ , we obtain

$$\frac{3\rho^2}{\beta(1 + \rho^2) \|x_k\|^2} < \mu_k < \frac{3\rho}{\beta(1 + \rho^2) \|x_k\|^2} \quad (84)$$

Note that,  $\arg \min_{\rho} \rho^2 / (1 + \rho^2) = 0$ , and  $\arg \max_{\rho} \rho / (1 + \rho^2) = 1$ . So making  $\rho = 0$  and  $\rho = 1$  in the lower and upper bounds, respectively, and replacing  $\|x_k\|^2$  with  $E[\|x_k\|^2]$ , we find the widest stochastic stability bound on  $\mu_k$  as follows:

$$0 < \mu < \frac{3}{2\beta E[\|x_k\|^2]}. \quad (85)$$

The significance of (85) is that it can easily be measured from the equalizer input samples. In adaptive filter theory, it is convenient to replace  $E[\|x_k\|^2]$  with  $\text{tr}(\mathbf{R})$ , where  $\mathbf{R} = E[x_k x_k^H]$  is the autocorrelation matrix of channel observation. Also note that, when noise is negligible and channel coefficients are normalized, the quantity  $\text{tr}(\mathbf{R})$  can be expressed as the product of equalizer length ( $N$ ) and transmitted signal average energy ( $P_a$ ); it gives

$$0 < \mu_{\beta\text{CMA}} < \frac{3}{2\beta N P_a} \quad (86)$$

Note that the bound (86) is remarkably similar to the stability bound of complex-valued LMS algorithm (refer to expression (20)). Comparing (86) and (20), we obtain a simple and elegant relation between the step-sizes of  $\beta\text{CMA}$  and complex-valued LMS:

$$\frac{\mu_{\beta\text{CMA}}}{\mu_{\text{LMS}}} < \frac{3}{\beta}. \quad (87)$$

## 6. Simulation results

We compare  $\beta\text{CMA}$  with CMA (Equation (27)) and SFA (Equation (40)). We consider transmission of amplitude-phase shift-keying (APSK) signals over a complex voice-band channel (channel-1), taken from Picchi & Prati (1987), and evaluate the average ISI traces at SNR = 30dB. We employ a seven-tap equalizer with central spike initialization and use 8- and 16-APSK signalling. Step-sizes have been selected such that all algorithms reached steady-state requiring same number of iterations. The parameter  $\beta$  is obtained as:

$$\beta = \begin{cases} 1.535, & \text{for 8-APSK} \\ 1.559, & \text{for 16-APSK} \end{cases} \quad (88)$$

Results are summarized in Fig. 6; note that the  $\beta\text{CMA}$  performed better than CMA and SFA by yielding much lower ISI floor. Also note that SFA performed slightly better than CMA.

Next we validate the stability bound (86). Here we consider a second complex channel (as channel-2) taken from Kennedy & Ding (1992). In all cases, the simulations were performed with  $N_{\text{it}} = 10^4$  iterations,  $N_{\text{run}} = 100$  runs, and no noise. In Fig. 7, we plot the probability of divergence ( $P_{\text{div}}$ ) for three different equalizer lengths, against the normalized step-size,  $\mu_{\text{norm}} = \mu / \mu_{\text{bound}}$ . The  $P_{\text{div}}$  is estimated as  $P_{\text{div}} = N_{\text{div}} / N_{\text{run}}$ , where  $N_{\text{div}}$  indicates the number of times equalizer diverged. Equalizers were initialized close to zero-forcing solution. It can be seen that the bound does guarantee a stable performance when  $\mu < \mu_{\text{bound}}$ .

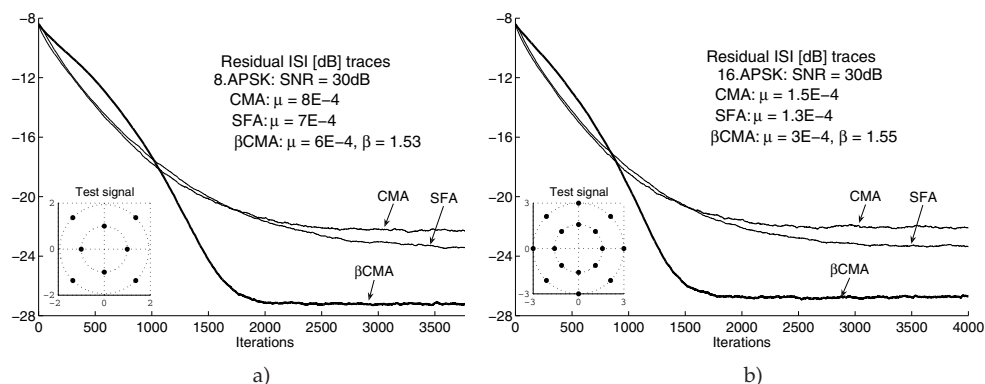


Fig. 6. Residual ISI: a) 8-APSK and b) 16-APSK. The inner and outer moduli of 8-APSK are 1.000 and 1.932, respectively. And the inner and outer moduli of 16-APSK are 1.586 and 3.000, respectively. The energies of 8-APSK and 16-APSK are 2.366 and 5.757, respectively.

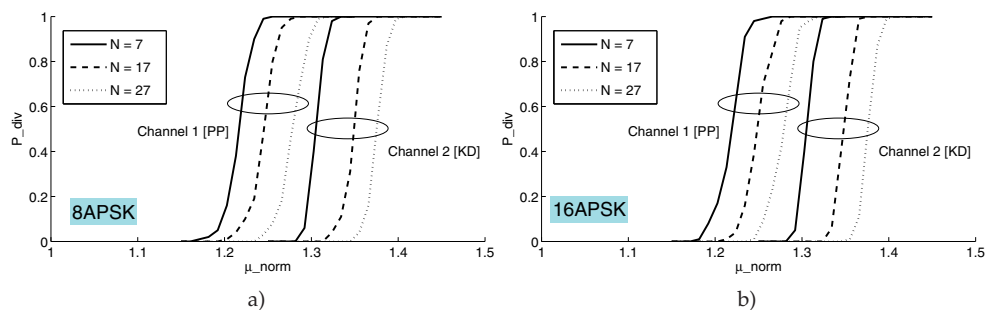


Fig. 7. Probability of divergence on channel-1 and channel-2 with three equalizer lengths, no noise,  $N_{it} = 10^4$  iterations and  $N_{run} = 100$  runs for a) 8-APSK and b) 16-APSK.

## 7. Concluding remarks

In this Chapter, we have introduced the basic concept of adaptive blind equalization in the context of single-input single-output communication systems. The key challenge of adaptive blind equalizers lies in the design of special cost functions whose minimization or maximization result in the removal of inter-symbol interference. We have briefly discussed popular criteria of equalization like Lucky, the mean square error, the minimum entropy, the constant modulus, Shalvi-Weinstein and Shtrom-Fan. Most importantly, based on minimum entropy deconvolution principle, the idea of designing specific cost function for the blind equalization of given transmitted signal is described in detail. We have presented a case study of amplitude-phase shift-keying signal for which a cost function is derived and corresponding adaptive algorithm is obtained. We have also addressed the admissibility of the proposed cost function and stability of the corresponding algorithm. The blind adaptation of the derived algorithm is shown to possess better convergence behavior compared to two existing algorithms. Finally, hints are provided to obtain blind equalization cost functions for square and cross quadrature amplitude modulation signals.

## 8. Exercises

1. Refer to Fig. 8 for geometrical details of square- and cross-QAM signals. Now following the ideas presented in Section 5, show that the blind equalization cost functions for square- and cross-QAM signals are respectively as follows:

$$\max_w E \left[ |y_n|^2 \right], \text{ s.t. } \max \{ |y_{R,n}|, |y_{I,n}| \} \leq R. \quad (89)$$

and

$$\max_w E \left[ |y_n|^2 \right], \text{ s.t. } \max \{ |\rho y_{R,n}|, |y_{I,n}| \} + \max \{ |y_{R,n}|, |\rho y_{I,n}| \} - \max \{ |y_{R,n}|, |y_{I,n}| \} \leq \rho R. \quad (90)$$

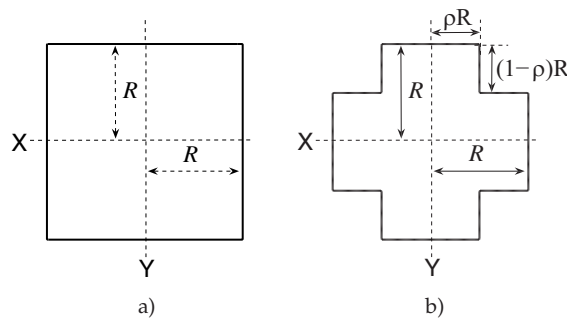


Fig. 8. Geometry of a) square- and b) cross-QAM  $\left(\rho = \frac{2}{3}\right)$ .

2. By exploiting the independence between the in-phase and quadrature components of square QAM signal, show that the following blind equalization cost function may be obtained:

$$\max_w E \left[ |y_n|^2 \right], \text{ s.t. } \max \{ |y_{R,n}| \} = \max \{ |y_{I,n}| \} \leq R. \quad (91)$$

The cost (91) originally appeared in Satorius & Mulligan (1993). Refer to Meng et al. (2009) and Abrar & Nandi (2010a), respectively, for its block-iterative and adaptive optimization.

## 9. Acknowledgment

The authors acknowledge the support of COMSATS Institute of Information Technology, Islamabad, Pakistan, King Fahd University of Petroleum and Minerals, Dhahran, Saudi Arabia, and the University of Liverpool, UK towards the accomplishment of this work.

## 10. References

- Abrar, S. & Nandi, A.K. (2010a). Adaptive solution for blind equalization and carrier-phase recovery of square-QAM, *IEEE Sig. Processing Lett.* 17(9): 791–794.
- Abrar, S. & Nandi, A.K. (2010b). Adaptive minimum entropy equalization algorithm, *IEEE Commun. Lett.* 14(10): 966–968.

- Abrar, S. & Nandi, A.K. (2010c). An adaptive constant modulus blind equalization algorithm and its stochastic stability analysis, *IEEE Sig. Processing Lett.* 17(1): 55–58.
- Abrar, S. & Nandi, A.K. (2010d). A blind equalization of square-QAM signals: a multimodulus approach, *IEEE Trans. Commun.* 58(6): 1674–1685.
- Abrar, S. & Shah, S. (2006a). New multimodulus blind equalization algorithm with relaxation, *IEEE Sig. Processing Lett.* 13(7): 425–428.
- Abrar, S. & Qureshi, I.M. (2006b). Blind equalization of cross-QAM signals, *IEEE Sig. Processing Lett.* 13(12): 745–748.
- Abrar, S., Zerguine, A. & Deriche, M. (2005). Soft constraint satisfaction multimodulus blind equalization algorithms, *IEEE Sig. Processing Lett.* 12(9): 637–640.
- Allen, J. & Mazo, J. (1973). Comparison of some cost functions for automatic equalization, *IEEE Trans. Commun.* 21(3): 233–237.
- Allen, J. & Mazo, J. (1974). A decision-free equalization scheme for minimum-phase channels, *IEEE Trans. Commun.* 22(10): 1732–1733.
- Baykal, B., Tanrikulu, O., Constantinides, A. & Chambers, J. (1999). A new family of blind adaptive equalization algorithms, *IEEE Sig. Processing Lett.* 3(4): 109–110.
- Bellini, S. (1986). Bussgang techniques for blind equalization, *Proc. IEEE GLOBECOM'86* pp. 1634–1640.
- Bellini, S. (1994). *Bussgang techniques for blind deconvolution and equalization*, in *Blind deconvolution*, S. Haykin (Ed.), Prentice Hall, pp. 8–59.
- Benedetto, F., Giunta, G. & Vandendorpe, L. (2008). A blind equalization algorithm based on minimization of normalized variance for DS/CDMA communications, *IEEE Trans. Veh. Tech.* 57(6): 3453–3461.
- Benveniste, A. & Goursat, M. (1984). Blind equalizers, *IEEE Trans. Commun.* 32(8): 871–883.
- Benveniste, A., Goursat, M. & Ruget, G. (1980a). Analysis of stochastic approximation schemes with discontinuous and dependent forcing terms with applications to data communications algorithms, *IEEE Trans. Automat. Contr.* 25(12): 1042–1058.
- Benveniste, A., Goursat, M. & Ruget, G. (1980b). Robust identification of a nonminimum phase system: Blind adjustment of a linear equalizer in data communication, *IEEE Trans. Automat. Contr.* 25(3): 385–399.
- Claerbout, J. (1977). Parsimonious deconvolution, *SEP-13*.
- C.R. Johnson, Jr., Schniter, P., Endres, T., Behm, J., Brown, D. & Casas, R. (1998). Blind equalization using the constant modulus criterion: A review, *Proc. IEEE* 86(10): 1927–1950.
- Ding, Z. & Li, Y. (2001). *Blind Equalization and Identification*, Marcel Dekker Inc., New York.
- Donoho, D. (1980). On minimum entropy deconvolution, *Proc. 2nd Applied Time Series Symp.* pp. 565–608.
- Farhang-Boroujeny, B. (1998). *Adaptive Filters*, John Wiley & Sons.
- Fiori, S. (2001). A contribution to (neuromorphic) blind deconvolution by flexible approximated Bayesian estimation, *Signal Processing* 81: 2131–2153.
- Garth, L., Yang, J. & Werner, J.-J. (1998). An introduction to blind equalization, *ETSI/STS Technical Committee TM6, Madrid, Spain* pp. TD7: 1–16.
- Godard, D. (1980). Self-recovering equalization and carrier tracking in two-dimensional data communications systems, *IEEE Trans. Commun.* 28(11): 1867–1875.
- Goupil, A. & Palicot, J. (2007). New algorithms for blind equalization: the constant norm algorithm family, *IEEE Trans. Sig. Processing* 55(4): 1436–1444.
- Gray, W. (1978). Variable norm deconvolution, *SEP-14*.



- Gray, W. (1979a). A theory for variable norm deconvolution, *SEP-15*.
- Gray, W. (1979b). *Variable Norm Deconvolution*, PhD thesis, Stanford Univ.
- Haykin, S. (1994). *Blind deconvolution*, Prentice Hall.
- Haykin, S. (1996). *Adaptive Filtering Theory*, Prentice-Hall.
- Im, G.-H., Park, C.-J. & Won, H.-C. (2001). A blind equalization with the sign algorithm for broadband access, *IEEE Commun. Lett.* 5(2): 70–72.
- Kennedy, R. & Ding, Z. (1992). Blind adaptive equalizers for quadrature amplitude modulated communication systems based on convex cost functions, *Opt. Eng.* 31(6): 1189–1199.
- Lucky, R. (1965). Automatic equalization for digital communication, *The Bell Systems Technical Journal* XLIV(4): 547–588.
- Lucky, R. (1966). Techniques for adaptive equalization of digital communication systems, *The Bell Systems Technical Journal* pp. 255–286.
- Meng, C., Tuqan, J. & Ding, Z. (2009). A quadratic programming approach to blind equalization and signal separation, *IEEE Trans. Sig. Processing* 57(6): 2232–2244.
- Nikias, C. & Petropulu, A. (1993). *Higher-order spectra analysis a nonlinear signal processing framework*, Englewood Cliffs, NJ: Prentice-Hall.
- Ooe, M. & Ulrych, T. (1979). Minimum entropy deconvolution with an exponential transformation, *Geophysical Prospecting* 27: 458–473.
- Picchi, G. & Prati, G. (1987). Blind equalization and carrier recovery using a ‘stop-and-go’ decision-directed algorithm, *IEEE Trans. Commun.* 35(9): 877–887.
- Pinchas, M. & Bobrovsky, B.Z. (2006). A maximum entropy approach for blind deconvolution, *Sig. Processing* 86(10): 2913–2931.
- Pinchas, M. & Bobrovsky, B.Z. (2007). A novel HOS approach for blind channel equalization, *IEEE Trans. Wireless Commun.* 6(3): 875–886.
- Romano, J., Attux, R., Cavalcante, C. & Suyama, R. (2011). *Unsupervised Signal Processing: Channel Equalization and Source Separation*, CRC Press Inc.
- Rupp, M. & Sayed, A.H. (2000). On the convergence analysis of blind adaptive equalizers for constant modulus signals, *IEEE Trans. Commun.* 48(5): 795–803.
- Sato, Y. (1975). A method of self-recovering equalization for multilevel amplitude modulation systems, *IEEE Trans. Commun.* 23(6): 679–682.
- Satorius, E. & Mulligan, J. (1992). Minimum entropy deconvolution and blind equalisation, *IEE Electronics Lett.* 28(16): 1534–1535.
- Satorius, E. & Mulligan, J. (1993). An alternative methodology for blind equalization, *Dig. Sig. Process.: A Review Jnl.* 3(3): 199–209.
- Serra, J. & Esteves, N. (1984). A blind equalization algorithm without decision, *Proc. IEEE ICASSP* 9(1): 475–478.
- Shalvi, O. & Weinstein, E. (1990). New criteria for blind equalization of non-minimum phase systems, *IEEE Trans. Inf. Theory* 36(2): 312–321.
- Shtrom, V. & Fan, H. (1998). New class of zero-forcing cost functions in blind equalization, *IEEE Trans. Signal Processing* 46(10): 2674.
- Sidak, Z., Sen, P. & Hajek, J. (1999). *Theory of Rank Tests*, Academic Press; 2/e.
- Treichler, J. & Agee, B. (1983). A new approach to multipath correction of constant modulus signals, *IEEE Trans. Acoust. Speech Signal Processing* 31(2): 459–471.
- Treichler, J. R. & Larimore, M. G. (1985). New processing techniques based on the constant modulus adaptive algorithm, *IEEE Trans. Acoust., Speech, Sig. Process.* ASSP-33(2): 420–431.

- Tugnait, J.K., Shalvi, O. & Weinstein, E. (1992). Comments on 'New criteria for blind deconvolution of nonminimum phase systems (channels)' [and reply], *IEEE Trans. Inf. Theory* 38(1): 210–213.
- Walden, A. (1985). Non-Gaussian reflectivity, entropy, and deconvolution, *Geophysics* 50(12): 2862–2888.
- Walden, A. (1988). A comparison of stochastic gradient and minimum entropy deconvolution algorithms, *Signal Processing* 15: 203–211.
- Wesolowski, K. (1987). Self-recovering adaptive equalization algorithms for digital radio and voiceband data modems, *Proc. European Conf. Circuit Theory and Design* pp. 19–24.
- Widrow, B. & Hoff, M.E. (1960). Adaptive switching circuits, *Proc. IRE WESCON Conf. Rec.* pp. 96–104.
- Widrow, B., McCool, J. & Ball, M. (1975). The complex LMS algorithm, *Proc. IEEE* 63(4): 719–720.
- Wiggins, R. (1977). Minimum entropy deconvolution, *Proc. Int. Symp. Computer Aided Seismic Analysis and Discrimination*.
- Wiggins, R. (1978). Minimum entropy deconvolution, *Geoexploration* 16: 21–35.
- Yang, J., Werner, J.-J. & Dumont, G. (2002). The multimodulus blind equalization and its generalized algorithms, *IEEE Jr. Sel. Areas Commun.* 20(5): 997–1015.
- Yuan, J.-T. & Lin, T.-C. (2010). Equalization and carrier phase recovery of CMA and MMA in blind adaptive receivers, *IEEE Trans. Sig. Processing* 58(6): 3206–3217.
- Yuan, J.-T. & Tsai, K.-D. (2005). Analysis of the multimodulus blind equalization algorithm in QAM communication systems, *IEEE Trans. Commun.* 53(9): 1427–1431.



## **Digital Communication**

Edited by Prof. C Palanisamy

ISBN 978-953-51-0215-1

Hard cover, 208 pages

**Publisher** InTech

**Published online** 07, March, 2012

**Published in print edition** March, 2012

All marketing is digital and everyone should have a digital strategy. Everything is going mobile. "The world has never been more social" is the recent talk in the community. Digital Communication is the key enabler of that. Digital information tends to be far more resistant to transmit and interpret errors than information symbolized in an analog medium. This accounts for the clarity of digitally-encoded telephone connections, compact audio disks, and much of the enthusiasm in the engineering community for digital communications technology. A contemporary and comprehensive coverage of the field of digital communication, this book explores modern digital communication techniques. The purpose of this book is to extend and update the knowledge of the reader in the dynamically changing field of digital communication.

### **How to reference**

In order to correctly reference this scholarly work, feel free to copy and paste the following:

Shafayat Abrar, Azzedine Zerguine and Asoke Kumar Nandi (2012). Adaptive Blind Channel Equalization, Digital Communication, Prof. C Palanisamy (Ed.), ISBN: 978-953-51-0215-1, InTech, Available from: <http://www.intechopen.com/books/digital-communication/adaptive-blind-channel-equalization>

**INTECH**  
open science | open minds

### **InTech Europe**

University Campus STeP Ri  
Slavka Krautzeka 83/A  
51000 Rijeka, Croatia  
Phone: +385 (51) 770 447  
Fax: +385 (51) 686 166  
[www.intechopen.com](http://www.intechopen.com)

### **InTech China**

Unit 405, Office Block, Hotel Equatorial Shanghai  
No.65, Yan An Road (West), Shanghai, 200040, China  
中国上海市延安西路65号上海国际贵都大饭店办公楼405单元  
Phone: +86-21-62489820  
Fax: +86-21-62489821

© 2012 The Author(s). Licensee IntechOpen. This is an open access article distributed under the terms of the [Creative Commons Attribution 3.0 License](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.