Blind Equalization

Christoph Krall
Graz University of Technology
Christian Doppler Laboratory for Nonlinear Signal Processing
Inffeldgasse 12, A-8010 Graz, Austria
christoph.krall@TUGraz.at

***Abstract*— Equalization is a very essential topic in data communication. Due to the communication channel, which can be wireline or wireless and furthermore time variant, the transmitted data symbols are distorted linearly over the channel. However, to overcome the effects of the channel, one can employ an equalizer which is adapted during the transmission of a training sequence. If transmission of training is not possible or not available, other ways have to be found to improve system performance.**
If the communication signals belong to a finite alphabet, signal properties are exploited to find an equalizer without transmission of training data. Generally, a cost function is defined which is iteratively minimized and an optimal solution is found. Other methods which try to estimate Higher Order Statistics of the signals and according to these estimates an inverse is computed. This paper shows a short overview about blind equalization strategies and tries to sketch a few basic ideas when designing algorithms for blind equalization.

I. Introduction

Equalization is a widely studied topic throughout all areas of communications, and signal processing. To achieve data reconstruction from distorted and/or noisy signal sources deconvolution has to be performed [1], [2] . Over the years many algorithms for *data aided* equalization have been proposed. These can be categorized in three different classes according to the criterion which is used to optimize the coefficients of the equalizer, i.e.,

- Zero-Forcing (ZF) Equalizer
- Minimum Mean Square Error (MMSE) Equalizer
- Minimum Bit Error Rate (MBER) Equalizer.

Conversely, equalization of distorted signals is also possible without the a priori knowledge of either the channel or training data. The adaptive equalizer is generally able to adapt to a well behaving channel inverse when restrictions on the statistical properties of the used signals are given. This is usually the case when we talk about communication signals since most communication systems employ signal constellations to transmit binary data [3], [4]. The equalizer has to adapt without any support from some system identification or parameter estimation, the process is often called *self recovering equalization*. The more generally term *Blind Equalization* was specified and embossed by *Beneviste* and *Gursat* in [5].

Generally, a communication channel can be represented by a filter as depicted in Fig. 1. The transmitted data symbols $\{s[k]\}$ belong to a finite alphabet $\mathcal{A}$, which can be defined as $\mathcal{A} = \{+1. -1\}, \{1+j, 1-j, -1+1, -1-j\}, \ldots$ typically visualized in a constellation diagram. The receiver has no information about the propagation channel $h(t)$ which is in our further discussions assumed to be linear and time invariant. However, in wireless communications the mobile radio propagation channel will have time variant behavior [1], [6] but is still linear. The output signal of the channel $x(t)$ may be additionally disturbed by noise $w(t)$ which is assumed to be i.i.d. Gaussian.

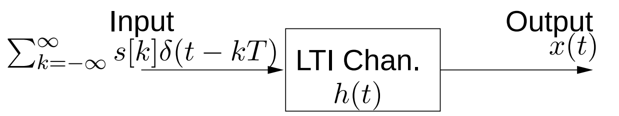

Fig. 1. Communication scenario

Considering the model for the communication channel, the objectives for the equalizer $G(z,\theta)$ are easily formulated. The combined channel and equalizer response in discrete-time domain should be just a scaled and delayed version of an ideal delta pulse, i.e.,

$$h[k] = a\delta[k-\Delta]. \tag{1}$$

This is often referred to as the *distortion-less* channel. Assuming the channel is known, analytic expressions for ZF and MMSE equalizers are easily given as [1]

$$G_{zf}(z,\theta) = \frac{z^{-\Delta}}{H(z)} \tag{2}$$

and

$$G_{mmse}(z,\theta) = \frac{H^*(z)z^{-\Delta}}{H(z)H^*(z^{-1}) + S_w(f)}, \tag{3}$$

where $H(z)$ represents the z-transform of the channel impulse response and $S_w(f)$ is the Power Spectral Density (PSD) of the noise process, respectively.

II. Blind Equalization

Since there is no training for parameter estimation available in blind equalization the used algorithms have to exploit signal properties. One specific property of linear systems is that the output PSD relates to the input PSD by

$$S_x(\omega) = S_s(\omega)|H(e^{j\omega})|^2 + S_w(\omega), \tag{4}$$

where $S_s(\omega)$ is the PSD of the source symbols $s[k]$. It is seen in (4) that a second order statistics measure as the PSD contains just magnitude information about the frequency response of the channel. Phase information is not available in this case and higher order moments/cumulants or spectra have to be considered when phase equalization has to be performed. One possible example is the to use Higher Order Statistics (HOS) to obtain the *trispectrum* of the output as

$$\begin{aligned} T_x(\omega_1, \omega_2, \omega_3) =& T_s(\omega_1, \omega_2, \omega_3) H(e^{j\omega_1}) H(e^{j\omega_2}) \\ & \times H(e^{j\omega_3}) H(e^{-j(\omega_1+\omega_2+\omega_3)}) \\ & + T_w(\omega_1, \omega_2, \omega_3) \end{aligned} \tag{5}$$

where for i.i.d. input signals $s[k]$ and Gaussian noise $w[k]$ the trispectra are constant, i.e.,

$$T_s(\omega_1, \omega_2, \omega_3) = \gamma_s, \quad T_w(\omega_1, \omega_2, \omega_3) = 0 \tag{6}$$

and the phase can be obtained as

$$\begin{aligned} \angle T_x(\omega_1, \omega_2, \omega_3) = & \angle H(e^{j\omega_1}) + \angle H(e^{j\omega_2}) + \angle H(e^{j\omega_3}) \\ & - \angle H(e^{-j(\omega_1+\omega_2+\omega_3)}) + \angle \gamma_s. \end{aligned} \tag{7}$$

As mentioned already, second order statistics (SOS) only provide magnitude information of the channel and all SOS methods are insufficient for blind equalization of a mixed phase channel containing zeros inside and outside the unit circle. Furthermore, it is not possible to identify a mixed phase channel from it's outputs if the input is i.i.d. Gaussian since only second order statistics are available [7]. Although the exact inverse of a non-minimum phase channel is unstable, a truncated anti-causal expansion can be delayed by Δ to allow causal FIR approximation of a ZF equalizer. As a further fact, ZF equalizers cannot be implemented for channels $H(z)$ with zeros on the unit circle. Any FIR approximation will have unbounded approximation error.

Generally we can distinguish between two different approaches when designing algorithms for blind equalization problems. First of all there is a huge class of Stochastic Gradient Descent (SGD) Algorithms which are suitable for equalizing communication channels blindly. The main idea behind the algorithms is that a cost function is minimized iteratively w.r.t. a parameter vector θ which represents the equalizer coefficients. Secondly, there are algorithms which try to exploit higher order statistical properties as we have seen before. These statistical informations are computed over a block of collected data and equalizers are computed according to equations (4), (5), and (7). Please note that generally all ways to find an equalizer $G(z, \theta)$ can be done in either time or frequency domain, thus we want to specify the parameter vector $\theta = [\theta_0 \; \theta_1 \; \dots \; \theta_m]^T$ and its corresponding z-transform as

$$G(z, \theta) = \sum_{i=0}^{m} \theta_i z^{-i} \tag{8}$$

Our objective to find a parameter set $\theta(k)$ at iteration step k such that the overall impulse response of the channel and the equalizer are a scaled and delayed versions of the Dirac delta pulse (*cf.* (1)). Furthermore, our discussion is restricted to real signals. However, please note that an extension to complex signals is straight forward. For that reason we restrict our alphabet $\mathcal{A}$ to symbols which meet

$$\{\pm(M-1)d, \, \pm(m-3)d, \, \dots, \, \pm 3d, \, \pm d\}, \quad M \text{ even}, \tag{9}$$

i.e., a M-level Pulse Amplitude Modulation (M-PAM). For the iterative approach we define a cost function which shall be a measure of the Inter-Symbol Interference (ISI). Thus, the *mean cost function* is given by

$$J(\theta) = E\{\Psi(y[k])\}, \tag{10}$$

where $\Psi(\cdot)$ is a scalar function of the equalizer output, which is preferably even to distinguish between $\pm$ levels. Generally we have to say that not necessarily a min. MSE equalizer is found anymore by optimizing the given cost function in (10). Merely the designed equalizer is a minimum ISI equalizer since we penalize ISI and try to optimize the coefficients to minimize ISI.

A stochastic gradient algorithm for an arbitrary cost function is given as

$$\begin{aligned} \theta(k+1) &= \theta(k) - \mu \frac{\partial}{\partial \theta(k)} \Psi(y[k]) \\ &= \theta(k) - \mu \Psi' \left(\theta(k)^T \mathbf{x}[k] \right) \mathbf{x}[k]. \end{aligned} \tag{11}$$

Thus we want to define the derivative of the *error function* Ψ as

$$\psi(x) = \Psi'(x) = \frac{\partial}{\partial x}\Psi(x), \tag{12}$$

which yields as a general *blind* update rule for LMS-type algorithms

$$\theta(k+1) = \theta(k) - \mu\psi\left(\theta(k)^T\mathbf{x}[k]\right)\mathbf{x}[k], \tag{13}$$

where μ is the stepsize of the algorithm.

III. Popular Algorithms for Blind Equalization

In the following we want to consider several popular algorithms for blind equalization. Most of them are based on a SGD search but each of them has a different cost/error function which usually improves the alorithm performance.

A. Sato Algorithm

Sato was the first to develop an algorithm for self recovering blind equalization in 1975 [8]. He considered a scenario for Binary Phase Shift Keyed (BPSK) modulated signals, i.e. $s[k] = \pm 1$. The error function is then given as

$$\psi(y[k]) = y[k] - R_1\text{sgn}(y[k]) \tag{14}$$

where R_1 is defined as

$$R_1 = \frac{E\{|s[k]|^2\}}{E\{|s[k]|\}}. \tag{15}$$

Thus considering an LMS-type adaptive algorithm, the parameter vector is updated as

$$\begin{aligned}\theta(k+1) &= \theta(k) - \psi(y[k])\mathbf{x}[k] \\ &= \theta(k) - [y[k] - R_1\text{sgn}(y[k])]\mathbf{x}[k],\end{aligned} \tag{16}$$

where $\mathbf{x}[\mathbf{k}]$ represents the vector of input signals x at iteration step k.

B. BGR Algorithms

The class of algorithms defined by Beneviste, Goursat, and Ruget (BGR) in [9] are similar to Satos algorithm. However, they allow for more complicated constellations, i.e. higher order modulation. Generally they are defined by a class of error functions given as

$$\psi(y[k]) = \tilde{\psi}(y[k]) - R_b\text{sgn}(y[k]) \tag{17}$$

where according to the modulation order b the constant R_b is defined as

$$R_b = \frac{E\{\tilde{\psi}(s[k])s[k]\}}{E\{|s[k]|\}}. \tag{18}$$

The function $\tilde{\psi}(\cdot)$ is desired to be an odd and twice differentiable function given by

$$\tilde{\psi}\text{"}(x) \geq 0, \forall x \geq 0 \tag{19}$$

The class of cost functions defined in (17) is much more gerneral than the one specified in Satos algorithm. However, for the special case of $\tilde{\psi}(x) = x$ the BGR algorithm is equivalent to Satos algorithm.

C. Stop-And-Go Algorithms

To avoid convergence to local minima in the cost function which may result in poor performance *Picchi* and *Prati* invented the stop-and-go adaptation of blind equalization algorithms [10]. Despite from the fact that more complexity is introduced by some stop-and-go methodology, the reliability of the adaptation to a highly performant inverse is improved. The main idea behind the whole algorithm is to use two different algorithms and compute each cost function separately. If both algorithms tend to have the correct sign for the gradient descent direction the blind equalizer is adapted. Otherwise the coefficients of the previous iteration are used for further equalization.

Consider an example where we have two given error functions denoted as $\psi_1(y)$ and $\psi_2(y)$, respectively. For this example the coefficients of an adaptive blind equalizer would be updated according to (20).

D. Constant Modulus Algorithms

Constant Modulus Algorithms (CMA) were invented independently by Godard [11] and Treichler [12] at the beginning of the 80ies. The main idea behind the algorithm is that all occurring signal which are not constant modulus are penalized. Conversely to the previously shown algorithms, CMA has the big advantage that carrier recovery and equalization can be done independently. Any occurring frequency offset Δf is just seen from the algorithm as a phase rotation. The CMA cost function is however insensitive to the phase and thus a constant phase offset does not affect the performance of the algorithm. Another advantage of CMA is that an analog implementation might be possible and thus application for analog modulation signals as FM or PM is also feasible.

The CMA is obtained by integrating Sato's error function, i.e.,

$$\Psi_1(y[k]) = 1/2(|y[k]| - R_1)^2. \tag{21}$$

Furthermore, the error function is generalized to

$$\Psi_q(y[k]) = 1/(2q)(|y[k]|^q - R_q)^2, \quad q = 1, 2, \ldots \tag{22}$$

$$\theta(k+1) = \begin{cases} \theta(k) - \mu\psi_1(y[k])\mathbf{x}[k], & \text{if } \mathrm{sgn}[\psi_1(y[k])] = \mathrm{sgn}[\psi_2(y[k])] \\ \theta(k), & \text{if } \mathrm{sgn}[\psi_1(y[k])] \neq \mathrm{sgn}[\psi_2(y[k])] \end{cases} \tag{20}$$

Which results in an update equation for the SGD-type Algorithm as

$$\theta(k+1) = \theta(k) - \mu(|y[k]|^q - R_q)|y[k]|^{q-2}y[k]\mathbf{x}[k] \tag{23}$$

and for the special case of $q = 2$ the algorithm is called "Constant Modulus Algorithm", i.e. the channel signal has a constant modulus $|s[k]|^2 = R_2$.

E. Shalvi and Weinstein Algorithm

As one example for a Higher Order Statistics (HOS) based algorithm we want to study the Algorithm proposed by *Shalvi* and *Weinstein* in 1990 [13]. The kurtosis of the output of a linear system is given by

$$K_y = E\{|y[k]|^4\} - 2(E\{|y[k]^2|\})^2 - |E\{y[k]^2\}|^2. \tag{24}$$

The algorithm now minimizes the absolute value of the kurtosis of the output, i.e. $|K_y|$ w.r.t. a constant power constraint, i.e.,

$$E\{|y[k]^2|\} = E\{|s[k]^2|\} \tag{25}$$

It can be shown [13] that if we assume that $s[k]$ is i.i.d. the power constraint of the output can be rewritten as

$$E\{|y[k]^2|\} = E\{|s[k]^2|\} \sum_{i=-\infty}^{\infty} |c[i]|^2, \tag{26}$$

where $c[i] = \theta * h[k]$ the convolution of the channel and the equalizer. For the kurtosis of the output we obtain

$$K_y = K_s \sum_i |c[i]|^4. \tag{27}$$

Now the overall impulse response $c[i]$ should be just a scaled and delayed version of a delta pulse. During the derivation it turns out that if the required constant power constraint in (25) is met, then $|K_y| = |K_s|$ is only possible when the combined impulse response $c[n] = \delta[n-\Delta]$, i.e., is the response of a distortion-less channel. Thus the Shalvi-Weinstein equalizer tries to maximize $\sum_i |c[i]|^4$ w.r.t. $\sum_i |c[i]|^2 = 1$.

IV. Summary

We have seen that blind equalization is not a trivial task since no data is available to provide the update algorithm with feedback on the estimation quality of the channel impulse response or its inverse. However, if we restrict the signals we transmit to a restricted range (alphabet) it is possible to exploit signal statistics to achieve self recovering equalization. Most frequently a cost function is defined and is minimized by a SGD algorithm. The other approach tries to exploit higher order statistical measures and according to the computed quantities an inverse is found. We have seen several different algorithms (Sato, BGR, Stop-and-Go, CMA, Shalvi-Weinstein) which work quite well in practice. Generally it is very hard to analyze convergence behavior of all these blind algorithms. Thus, research is still ongoing in defining new cost functions (error functions) or HOS based approaches for various application scenarios.

References

[1] J. G. Proakis, *Digital Communications*, 4th ed. McGraw-Hill International, 2001.

[2] J. R. Barry, E. A. Lee, and D. G. Messerschmitt, *Digital Communications*, 3rd ed. Kluwer Academic Publishers, 2004.

[3] S. Haykin, Ed., *Blind Deconvolution*. PTR Prentice Hall, Inc., 1994.

[4] Z. Ding and Y. G. Li, *Blind Equalization and Identification*, ser. Signal Processing and Communication Series. Marcel Dekker, Inc., 2001.

[5] A. Beneviste and M. Goursat, "Blind Equalizers," *IEEE Trans. Commun.*, vol. COM-32, no. 8, pp. 871 – 883, August 1984.

[6] T. S. Rappaport, *Wireless Communications*, 2nd ed. Prentice Hall, 2002.

[7] A. Papoulis and S. U. Pillai, *Probability, Random Variables and Stochastic Processes*, 4th ed. McGraw-Hill International, 2002.

[8] Y. Sato, "A Method of Self-Recovering Equalization for Multilevel Amplitude-Modulation Systems," *IEEE Trans. Commun.*, vol. COM-23, pp. 679 – 682, June 1975.

[9] A. Beneviste, M. Goursat, and G. Ruget, "Robust Identification of a Nonminimum Phase System: Blind Adjustment of a Linear Equalizer in Data Communications," *IEEE Trans. Automat. Contr.*, vol. AC-25, no. 3, pp. 385 – 399, June 1980.

[10] G. Picchi and G. Prati, "Blind Equalization and Carrier Recovery Using a 'stop-and-go' decision-directed algorithm," *IEEE Trans. Commun.*, vol. COM-35, pp. 877 – 887, September 1987.

[11] D. N. Godard, "Self-Recovering Equalization and Carrier Tracking in Two-Dimensional Data Communication Systems," *IEEE Trans. Commun.*, vol. COM-28, pp. 1867 – 1875, November 1980.

[12] J. R. Treichler and M. G. Agee, "A New Approach to Multipath Correction of Constant Modulus Signals," *IEEE Trans. Acoust., Speech, Signal Processing*, no. ASSP-31, pp. 349 – 472, April 1983.

[13] O. Shalvi and E. Weinstein, "New Criteria for Blind Deconvolution of Nonminimum Phase Systems (Channels)," *IEEE Trans. Inform. Theory*, vol. 36, no. 2, pp. 312 – 321, March 1990.