A FAST BLIND EQUALIZATION METHOD BASED ON SUBGRADIENT PROJECTIONS

Can Kizilkale and Alper T. Erdogan

EE Department, Koc University
Sariyer, Istanbul, Turkey.
email:ckizilkale@ku.edu.tr
alperdogan@ku.edu.tr

ABSTRACT

A novel blind equalization method based on a subgradient search over a convex cost surface is proposed. This is an alternative to the existing iterative blind equalization approaches such as the Constant Modulus Algorithm (CMA) which mostly suffer from the convergence problems caused by their non-convex cost functions. The proposed method is an iterative algorithm, for both real and complex constellations, with a very simple update rule that minimizes the l_∞ norm of the equalizer output under a linear constraint on the equalizer coefficients. The algorithm has a nice convergence behavior attributed to the convex l_∞ cost surface. Examples are provided to illustrate the algorithm's performance.

1. INTRODUCTION

The blind equalization has been a research focus for several decades. The goal has been the development of fast, low complexity and robust algorithms which avoid the consumption of useful bandwidth by the training data.

Among the existing methods probably the most popular ones for practical applications are Constant Modulus (CM) [1] based algorithms due to their low-complexity implementation. On the other hand, the CM based algorithms have the disadvantage of ill or slow convergence due to the topology of the corresponding non-convex cost surfaces.

The convergence problems caused by the existence of local minima and saddle points can be solved by replacing the non-convex cost functions (such as CM and Maximum Likelihood) with convex cost functions. The reference [2] investigates the convex cost functions for blind equalization and proposes minimizing the l_∞ norm of the equalizer output under a linear constraint on the equalizer coefficient vector as a possible choice for the convex cost function. The choice of the l_∞ cost function solves the slow and ill convergence issues related to local minima and saddle points. However, in order to make the l_∞ norm minimization as a practical choice for blind equalization, the l_∞ norm minimization algorithms with low complexity need to be developed. In fact, this is the focus of our article: although the l_∞ cost function is not differentiable and not suitable for ordinary gradient search type iterative algorithm, we propose an iterative algorithm based on subgradient optimization.

The organization of the article is as follows: The blind equalization setup used throughout the article is provided in Section 2. Section 3 provides the background on l_∞ norm minimization as the cost function for blind equalization and Section 4 outlines the existing methods for subgradient optimization. Section 5 is where the application of subgradient methods to the blind equalization problem and the corresponding algorithms are provided. Finally, Section 6 provides examples for illustrating the performance of the algorithm, which is followed by conclusion.

2. BLIND EQUALIZATION SETUP

Throughout the paper, we assume the sample spaced equalization setup shown in Figure 1. Here $\{x_k \in \{-2\cdot M+1,...,2\cdot M-1\}\}$ (PAM) or $\{x_k = a_k + jb_k$ where $a_k, b_k \in \{-2\cdot M+1,...,2\cdot M-1\}\}$ (QAM) is the information sequence sent by the transmitter, $\{h_k; k \in \{0,...,N_h-1\}\}$ is the impulse response of the sample spaced channel, $\{y_k\}$ is the receiver input signal, $\{w_k\ k \in \{0,...,N_w-1\}\}$ is the equalizer and $\{z_k\}$ is the equalizer output.

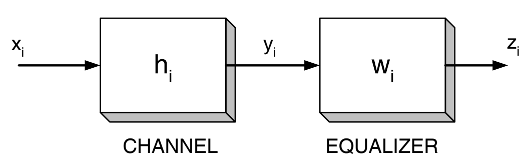

Fig. 1. The Equalization Setup

The purpose of the blind equalization is to make the equalizer output $\{z_k\}$ as close as possible to the delayed version of the input $\{x_k\}$, using only $\{y_k\}$ without any training data or a priori knowledge of the channel $\{h_k\}$.

3. l_∞ NORM AS THE CONVEX COST FUNCTION FOR BLIND EQUALIZATION

The reference [2] investigates the properties of convex functions as the candidates for the cost functions for the blind equalization for PAM signals. Considering the equalization setup in Figure 1, under the assumptions that the input constellation has the maximum magnitude symmetry around zero such that $\max x_k = (2 \cdot M - 1)$ and $\min x_k = -2 \cdot M + 1$, and $\{x_k\}$ is sufficiently rich in terms of variations in time, it can be easily shown that

$$||z||_\infty \quad = \quad (2 \cdot M - 1)||c||_1 \qquad (1)$$

where $c = h * w$ is the impulse response of the combined channel and equalizer. Therefore, minimizing the l_∞ norm of the output is equivalent to minimizing the l_1 norm of the overall impulse response. Under these assumptions, in reference [2], it is shown that the equalizer coefficients ($\{w_k; k \in \mathcal{Z}\}$) obtained by solving the following problem

$$\begin{aligned} &\text{minimize } ||z||_\infty && \textit{(Problem 1)} \\ &\text{s. t. } w_L = 1 \end{aligned}$$

for some $L \in \mathcal{Z}$ will be the scaled and time shifted inverse of $\{h_k\}$ such that

$$h * w = G\delta_{n-d} \qquad (2)$$

for some magnitude G and delay d. Placing the FIR constraint on equalizer coefficients will still preserve the convex nature of the problem. However, there will be a performance degradation due to this constraint.

What remains is to develop preferably low-complexity algorithms to solve the convex optimization problem. In [2], it is proposed to minimize p-norm of the output (with a large p value) to approximate the solution of *Problem 1*. In this article, we provide an algorithm to minimize the infinity norm directly.

For the solution of *Problem 1*, the real time implementation requirement places a limit on the time span for the maximum value search. In order to reflect this requirement into our algorithm, we can modify the convex optimization problem as

$$\begin{aligned} &\text{minimize } ||z \cdot r||_\infty && \textit{(Problem 2)} \\ &\text{s. t. } w_L = 1 \end{aligned}$$

where $\{r_n\}$ is a rectangular window function with size Ω. Hence the final problem would be equivalent to minimizing the l_∞ norm of the finite size vector

$$\underbrace{\begin{bmatrix} z_0 \\ z_1 \\ \vdots \\ z_{\Omega-1} \end{bmatrix}}_{\mathbf{z}} = \underbrace{\begin{bmatrix} y_0 & y_{-1} & \cdots & y_{-N_w+1} \\ y_1 & y_0 & \cdots & y_{-N_w+2} \\ \vdots & \ddots & \ddots & \vdots \\ y_{\Omega-1} & y_{\Omega-2} & \cdots & y_{\Omega-N_w} \end{bmatrix}}_{\mathbf{Y}} \underbrace{\begin{bmatrix} w_0 \\ w_1 \\ \vdots \\ w_{N_w-1} \end{bmatrix}}_{\mathbf{w}} \qquad (3)$$

Since $\mathbf{z}$ is a linear function of the search vector $\mathbf{w}$ and the constraint $w_L = 1$ is a linear constraint, the corresponding l_∞ norm minimization problem can be solved with the well-known Linear Programming (LP) formulation:

$$\begin{aligned} &\text{minimize } t && \textit{(Problem 3)} \\ &\text{s.t. } \begin{bmatrix} -\mathbf{1} & \mathbf{\Gamma} \\ -\mathbf{1} & -\mathbf{\Gamma} \end{bmatrix} \begin{bmatrix} t \\ \mathbf{w_s} \end{bmatrix} \leq \begin{bmatrix} -\mathbf{q} \\ \mathbf{q} \end{bmatrix} \end{aligned}$$

where Γ is equivalent to the Y matrix with $(L+1)^{st}$ column deleted, $\mathbf{q}$ is the $(L+1)^{st}$ column of Y, $\mathbf{w}_s$ is the $\mathbf{w}$ with $(L+1)^{st}$ element deleted (since $w_L = 1$), and $\begin{bmatrix} t & \mathbf{w_s}^T \end{bmatrix}^T$ is the search vector. There are various approaches to solve the LP of *Problem 3*. However, instead of LP approach, we propose a low complexity iterative method, which is more suitable for real time applications, that is based on the recently developed subgradient optimization methods.

In the next section, we provide a review of subgradient projection approaches in relevance to our discussion. Later we present the algorithms based on subgradient projections.

4. A REVIEW OF SUBGRADIENT METHODS

Let $f(\mathbf{w})$ be a convex and possibly non-differentiable function with domain S, where S is convex. The subdifferential of $f(\mathbf{w})$ at point $\mathbf{w}$ is defined as

$$\partial f(\mathbf{w}) = \{\mathbf{g} | f(\mathbf{y}) \geq f(\mathbf{w}) + \langle \mathbf{g}, \mathbf{y} - \mathbf{w} \rangle \quad \forall \mathbf{y} \in S\}, \quad (4)$$

where $\langle \cdot, \cdot \rangle$ is the inner product. A vector $\mathbf{g}$ which is a member of $\partial f(\mathbf{w})$ is called a subgradient of $f(\mathbf{w})$ at $\mathbf{w}$. The non-differentiable counterpart of the gradient-descent algorithm is the subgradient projection method in which the gradient is simply replaced by a subgradient:

$$\mathbf{w}^{(i+1)} = \mathcal{P}_S \left\{ \mathbf{w}^{(i)} - \mu^{(i)} \mathbf{g}^{(i)} \right\} \qquad (5)$$

where $\mathbf{g}^{(i)}$ is a subgradient picked from the subdifferential set $\partial f(\mathbf{w}^{(i)})$ and $\mathcal{P}_S$ is the projection to convex set S. Although the subgradient algorithm looks very much like the gradient descent algorithm, in the subgradient iteration it may happen that $f(\mathbf{w}^{(i+1)}) > f(\mathbf{w}^{(i)})$ for any $\mu^{(i)} > 0$[3]. However, if the $\mu^{(i)}$ parameter is properly chosen, $\mathbf{w}^{(i)}$ can be made to converge to the optimal point.

One major result about the selection of the step size parameter μ_i is due to Polyak [4]: if

$$\lim_{i \to \infty} \frac{\mu^{(i)}}{||\mathbf{g}^{(i)}||} = 0 \text{ and } \sum_{i=0}^{\infty} \frac{\mu^{(i)}}{||\mathbf{g}^{(i)}||} = \infty$$

hold then $\lim_{i \to \infty} \mathbf{w}^{(i)} = \mathbf{w}^*$, the optimal point, which provides sufficient conditions for the convergence.

Furthermore, if $\mathbf{w}^{(i)}$ is not the optimal point and if the step size satisfies

$$0 < \mu^{(i)} < 2\frac{(f(\mathbf{w}^{(i)}) - f^*)}{||\mathbf{g}^{(i)}||_2^2} \tag{6}$$

where f^* is the minimum value of $f(\mathbf{w})$, then it is guaranteed that

$$||\mathbf{w}^{(i+1)} - \mathbf{w}^*||_2 < ||\mathbf{w}^{(i)} - \mathbf{w}^*||_2 \qquad \forall i \tag{7}$$

i.e., the distance to the optimal vector decreases monotonically. As f^* is not known a priori in many practical problems, the use of an estimate of f^*, instead of f^* have been investigated in several references(see for example [5]. Recently Goffin and Kiwiel [6] and Sherali et. al. [7] also proposed simple and converging subgradient algorithms with variable target value $\hat{f}^*$.

5. ITERATIVE BLIND EQUALIZATION WITH SUBGRADIENT PROJECTIONS

In order to produce low-complexity iterative algorithms for solving the *Problem 2*, we apply the subgradient projection approach which is outlined in the previous section. The subdifferential set for the blind equalization cost function

$$f(\mathbf{w}) = \|\mathbf{z}\|_\infty = \|\mathbf{\Gamma}\mathbf{w}_s + \mathbf{q}\|_\infty \tag{8}$$

is given by

$$\begin{aligned} \partial f(\mathbf{w}_s) \quad = \quad & \mathbf{Co}\left\{\{\mathbf{\Gamma}_{k,:}^T | z_k = ||z||_\infty\}\cup\right. \\ & \left.\{-\mathbf{\Gamma}_{k,:}^T | z_k = -||z||_\infty\}\right\}, \end{aligned} \tag{9}$$

where $\mathbf{Co}\{\}$ is the convex-hull operation. Upon the inspection of the subdifferential set 9, the search direction for the subgradient projection algorithm are obtained from the equalizer input vectors causing the maximum magnitude equalizer output within the given window. If J is the set of time instants for which maximum magnitude is achieved, i.e., $J = \{k| \quad |z_k| = ||\mathbf{z}||_\infty\}$, then a possible search direction for the subgradient projection algorithm is

$$\mathbf{d} = -\sum_{k\in J} \xi_k \mathrm{sign}(z_k)\mathbf{\Gamma}_{k,:}^T \tag{10}$$

where $\sum_{k\in J}\xi_k = 1$ and $\xi_k \geq 0$. For convenience, one may choose $\xi_l = 1$ for some $l \in J$ and $\xi_k = 0$ for $k \neq l$ in which case the search direction simplifies to

$$\mathbf{d} = -\mathrm{sign}(z_l)\mathbf{\Gamma}_{l,:}^T. \tag{11}$$

As a result, we can write the update rule for the subgradient based blind equalization algorithm (SGBA) as

$$\mathbf{w}_s^{(i+1)} = \mathbf{w}_s^{(i)} - \mu^{(i)}\mathrm{sign}(z_{l^{(i)}}^{(i)})\mathbf{\Gamma}_{l^{(i)},:}^T, \tag{12}$$

where

- $l^{(i)} \in \{0, ..., \Omega - 1\}$ is the index where maximum magnitude output is achieved at the i^{th} iteration.
- $\mu^{(i)}$ is the step size at the i^{th} iteration. We suggest the use of
$$\mu^{(i)} = \alpha\frac{z_{l^{(i)}}^{(i)} - \hat{f^*}^{(i)}}{||\mathbf{\Gamma}_{l^{(i)},:}||_2^2}, \tag{13}$$
as in the relaxation rule of Equation 6 with $\alpha \in [0, 2)$. Here a reasonable choice for $\hat{f^*}^{(i)}$ is given by
$$\hat{f^*}^{(i)} = \frac{1}{\Omega}||\mathbf{\Gamma}\mathbf{w}_s^{(i)} + \mathbf{q}||_1, \tag{14}$$
which is the average of the magnitude of the output within the selected window. Note that for this choice, $\mu^{(i)}$ is nonnegative and it is equal to zero only under the perfect equalization condition. Alternatively, one could use the methods suggested in references [6, 7] to determine $\hat{f^*}^{(i)}$.

5.1. Complex Constellation Case

As shown in [8], the results of [2] for real constellations can be extended to the complex constellations that satisfy the property $\max_k |\Re e\{x_k\}| = \max_k |\mathcal{I}m\{x_k\}|$, where the corresponding optimization problem is defined as

$$\begin{aligned} &\text{minimize } \|\Re e\{\mathbf{z}\}\|_\infty = \|\Re e\{\mathbf{Y}\mathbf{w}\}\|_\infty \quad \textit{(Problem 4)} \\ &\text{s. t. } \Re e\{w_L\} = 1, \end{aligned}$$

where Y and $\mathbf{w}$ are as defined in Equation 3. Therefore, the update rule of the algorithm in this case becomes

$$\begin{aligned} \mathbf{w}^{(i+1)} &= \mathbf{w}^{(i)} - \mu^{(i)}\mathrm{sign}(\Re e\{z_{l^{(i)}}^{(i)}\})\mathbf{Y}_{l^{(i)},:}^H \\ \Re e\{w_L^{(i+1)}\} &= 1. \end{aligned}$$

The second equation in the update rule above is the projection to the constraint set.

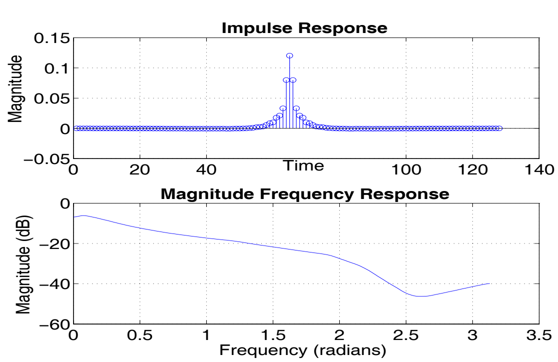

Fig. 2. DSL Channel.

6. EXAMPLES

In the first example, we consider a DSL channel for the ITU G.SHDSL Central Office Transmitter and the CSA4 Channel with 2 bridge taps. The corresponding 128 tap impulse response and its frequency response are shown in Figure 2. In this simulation, we assumed $M = 2, \Omega = 800, N_w = 31$, $L = 16$ and used the relaxation step rule provided in Equations 13 and 14. Figure 3 shows the corresponding open eye measure of the equalizer output as a function of iterations where open eye measure is defined as

$$\rho(c) = \frac{\sum |c_i| - p}{p} \tag{15}$$

where $c = h * w$ is the cascade of equalizer and channel and $p = \max_i |c_i|$ is its maximum magnitude tap.

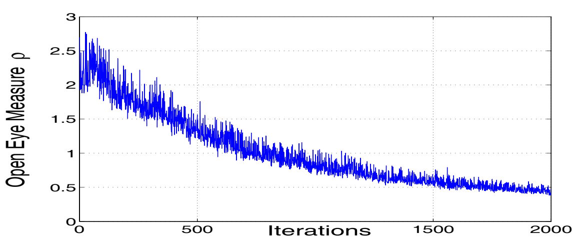

Fig. 3. Open-eye measure for the DSL Channel.

As another example, we consider the complex channel $h = \{-1.0493 + 0.2305i, 1.4129 - 1.4497i, -0.2540 + 0.2021i, 0.5302 - 0.7732i\}$ from the reference [9] with a 4-QAM input and an equalizer with length 21. The figure 4 provides the comparison of our algorithm with CMA, where each curve corresponds to a different initialization point(at same distance to the optimal point). As it can be seen from this figure, the subgradient based algorithm converges at less number of iterations and is less sensitive to the choice of initial point for the iterations due to convex nature of its cost function.

7. CONCLUSION

We introduced a novel iterative blind algorithm with a very simple update rule based on the minimization of l_∞ norm of the output of the equalizer using subgradient iterations. Due to the convex nature of the l_∞ cost function, the convergence problems that exist in CM type algorithms caused by the local extremum points and arbitrary initialization points don't exist. Updates are very simple and fast especially for DSP systems where equalizer is implemented in hardware. Finally, the iterative algorithm presented in this paper is suitable for sample space (an implicit Higher Order Statistics method) as well as fractionally spaced channels.

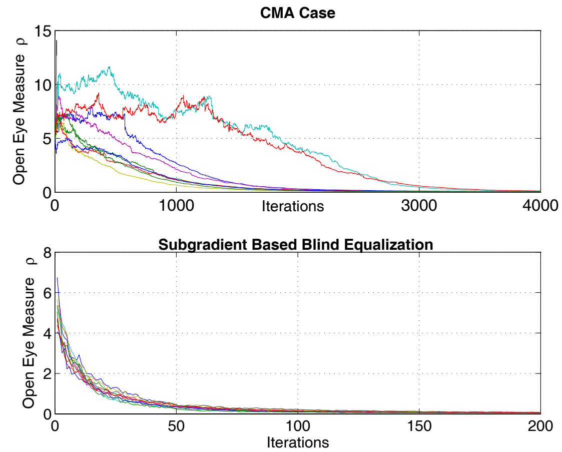

Fig. 4. Complex Channel:CMA vs. SGBA

8. REFERENCES

[1] C. Johnson, J. Schniter, T. Endres, J. Behm, R. Casas, V. Brown, and C. Berg, "Blind equalization using the constant modulus criterion: A review," *Proc. of IEEE, Special Issue on Blind System Identificaiton and Estimation*, pp. 1927–1950, 1998.

[2] S. Vembu, S. Verdu, R. Kennedy, and W. Sethares, "Convex cost functions in blind equalization," *IEEE TSP*, vol. 42, pp. 1952–1960, 1994.

[3] D. P. Bertsekas, *Nonlinear Programming*, Athena Scientific, 1999.

[4] B. T. Polyak, "A general method for solving extremal problems," *Doklady Akademii Nauk SSSR*, vol. 174, pp. 33–36, 1967.

[5] M. Bazaraa and H. Sherali, "On the choice of step size in subgradient optimization," *European Journal of OR.*, vol. 7, pp. 380–388, 1981.

[6] J. Goffin and K. Kiwiel, "Convergence of a simple subgradient level method," *Math. Programming*, vol. 85, pp. 207–211, 1999.

[7] H. Sherali, G. Choi, and C. Tuncbilek, "A variable target method for nondifferentiable optimization," *OR. Letters*, vol. 26, pp. 1–8, 2000.

[8] Z. Ding and Z. Luo, "A fast linear programming algorithm for blind equalization," *IEEE TCOM.*, vol. 48, pp. 1432–1436, 2000.

[9] C. Papadias, *Methods for Blind Equalization and Identification of Linear Channels*, Ph.D. Thesis, 1995.