

On the convergence of subgradient based blind equalization algorithm[☆]

Alper T. Erdogan^{*}

Department of Electrical and Electronics Engineering, Koc University, Rumelifeneri Yolu, 34450, Istanbul, Turkey

Received 6 July 2005; received in revised form 7 December 2005; accepted 12 March 2006

Available online 6 April 2006

Abstract

SubGradient based Blind Algorithm (SGBA) has recently been introduced [A.T. Erdogan, C. Kizilkale, Fast and low complexity blind equalization via subgradient projections, *IEEE Trans. Signal Process.* 53 (2005) 2513–2524; C. Kizilkale, A.T. Erdogan, A fast blind equalization method based on subgradient projections, *Proceedings of IEEE ICASSP 2004, Montreal, Canada, vol. 4, pp. 873–876.*] as a convex and low complexity approach for the equalization of communications channels. In this article, we analyze the convergence behavior of the SGBA algorithm for the case where the relaxation rule is used for the step size. Our analysis shows that the monotonic convergence curve for the mean square distance to the optimal point is bounded between two geometric-series curves, and the convergence rate is dependent on the eigenvalues of the correlation matrix of channel outputs. We also provide some simulation examples for the verification of our analytical results related to the convergence behavior.

© 2006 Elsevier B.V. All rights reserved.

Keywords: Blind equalization; Convergence analysis; Convex optimization; SGBA; Subgradient

1. Introduction

Blind equalization has been an active research field where various algorithms have been proposed. All of these algorithms exploit pieces of information available to recover an original source signal from its linearly distorted version, where the distortion process is unknown.

In the design of blind algorithms for equalization we can name two important criteria: convergence

behavior and implementation complexity. Convergence behavior is critical as many algorithms suffer from ill or slow convergence (due to the structure of their corresponding non-convex cost surfaces [1,2]). The convergence issue has been addressed in [3], where the use of convex cost functions was proposed as a viable solution to the convergence problem.

Implementation complexity is also critical as these algorithms have target applications that include real-time communications systems with high symbol rates. Therefore, for such systems, a fast convergence behavior with minimal computational and hardware resource consumption is very desirable.

In the field of blind equalization with convex cost functions, linear programming based algorithms [4,5] have been proposed. These algorithms solve an

[☆]This work was supported in part by TUBITAK Career Award, Contract no: 104E073.

^{*}Tel.: +90 212 338 1490; fax: +90 212 338 1548.

E-mail address: alperdogan@ku.edu.tr.

URL: <http://aspc.ku.edu.tr>.

l_∞ norm based optimization problem, which is proposed as the convex approach in [3], by casting it as a linear programming problem, for both symbol spaced and fractionally spaced channels. However, their computational complexities are fairly high.

Recently, the SGBA (SubGradient based Blind Algorithm) algorithm has been introduced [6,7] as an alternative convex blind equalization method. The SGBA algorithm is an iterative method with a very simple update rule and it has a much lower computational complexity than the linear programming based convex blind approaches. Since the SGBA algorithm is based on a convex cost surface it does not suffer from the slow convergence problems caused by attraction to saddle points (although the slow convergence due to some other factors such as near flat regions in the cost function is still a possibility) and the ill-convergence problems due to the existence of false minima. Furthermore, the use of relaxation rule in the SGBA algorithm guarantees convergence with a monotonic decrease in the distance to the (global) optimal point.

The focus of this article is the analysis of the convergence of the SGBA algorithm using the relaxation step size rule. For this purpose, we first introduce the blind equalization setup and the SGBA algorithm in Section 2. In Section 3, we provide a geometric picture for the convergence of the SGBA algorithm. The mean square convergence analysis of SGBA is given in Section 4. Section 5 provides some simulation examples related to convergence behavior. Section 6 is the Conclusion.

2. Blind equalization setup and SGBA algorithm

The equalization setup that we use throughout the article is shown in Fig. 1. Here

- $\{x_i \in \{-2 \cdot M + 1, \dots, 2 \cdot M - 1\}\}$ is the information sequence sent by the transmitter,
- $\{h_i; i \in \{0, \dots, N_h - 1\}\}$ is the impulse response of the symbol-spaced channel which is the combination of the linear distortion caused by the

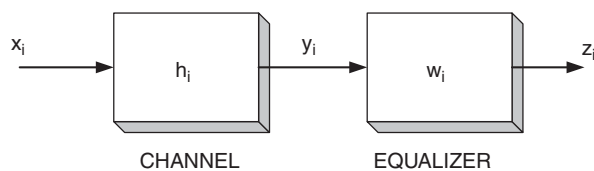


Fig. 1. The setup for blind equalization.

communication medium and the pulse shaping filter,

- $\{y_i\}$ is the input sequence to the equalizer,
- $\{w_i; i = 0, \dots, N_w - 1\}$ is the set of equalizer coefficients, and $\mathbf{w} = [w_0 \ w_1 \ \dots \ w_{N_w-1}]^T$ is the equalizer coefficient vector,
- $\{z_i\}$ is the equalizer output.

The above setup is for real PAM signals, however, both the setup and the corresponding convergence analysis can easily be generalized for complex constellations.

The convex optimization problem, on which the SGBA algorithm is based, is the minimization of the infinity-norm of equalizer output under a linear constraint on equalizer coefficients (such as a fixed tap constraint) [3]. We can write the corresponding optimization setting as

$$\begin{aligned} \min \quad & \|z \odot r\|_\infty \\ \text{s.t.} \quad & w_L = 1, \end{aligned}$$

where w_L is the fixed tap, $\{r_n\}$ is a rectangular window function, with

$$r_n = \begin{cases} 1 & 0 \leq n \leq \Omega - 1, \\ 0 & \text{otherwise} \end{cases} \quad (1)$$

and $z \odot r$ represents the element by element multiplication of two discrete time sequences (i.e., $(z \odot r)_k = z_k r_k$). This problem is equivalent to the minimization of the l_∞ norm of the finite size vector

$$\begin{aligned} \underbrace{\begin{bmatrix} z_0 \\ z_1 \\ \vdots \\ z_{\Omega-1} \end{bmatrix}}_z &= \underbrace{\begin{bmatrix} y_0 & y_{-1} & \cdots & y_{-N_w+1} \\ y_1 & y_0 & \cdots & y_{-N_w+2} \\ \vdots & \ddots & \ddots & \vdots \\ y_{\Omega-1} & y_{\Omega-2} & \cdots & y_{\Omega-N_w} \end{bmatrix}}_Y \underbrace{\begin{bmatrix} w_0 \\ w_1 \\ \vdots \\ w_{N_w-1} \end{bmatrix}}_w \\ &= \Gamma \mathbf{w}_s + \mathbf{q}, \end{aligned} \quad (2)$$

where Γ is equivalent to Y matrix with $(L+1)$ th column deleted, \mathbf{q} is the $(L+1)$ th column of Y , and \mathbf{w}_s is \mathbf{w} with $(L+1)$ th element deleted (since $w_L = 1$).

The SGBA algorithm to solve this optimization problem is given by the following iterations [7,6]:

$$\mathbf{w}_s^{(i+1)} = \mathbf{w}_s^{(i)} - \underbrace{\mu^{(i)} \text{sign}(z_{l^{(i)}}^{(i)}) \Gamma_{l^{(i)},:}^T}_{\mathbf{g}^{(i)}}, \quad (3)$$

where

- $l^{(i)} \in \{0, \dots, \Omega - 1\}$ 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 will assume that the relaxation step size rule is used where

$$\mu^{(i)} = \alpha \frac{|z_{l^{(i)}}^{(i)}| - f_{\text{opt}}}{\|\Gamma_{l^{(i)},:}\|_2^2}, \quad (4)$$

where f_{opt} is the optimal value of the objective function. Since f_{opt} is not known a priori, the step size rule in (4) may appear to be unreasonable. However, in [6], it is illustrated that through use of variable target schemes (which does not require the knowledge of f_{opt}), we can achieve similar convergence performance as the original relaxation rule. Therefore, we will use the ideal relaxation rule in our analysis.

3. A geometric picture for convergence

Based on the update equation in (3), we can write a recursion for the distance to the optimal point in the form:

$$\|\mathbf{w}^{(i+1)} - \mathbf{w}_{\text{opt}}\|^2 = \|\mathbf{w}_s^{(i+1)} - \mathbf{w}_{s,\text{opt}}\|^2 \quad (5)$$

$$= \|\mathbf{w}_s^{(i)} - \mu^{(i)} \mathbf{g}^{(i)} - \mathbf{w}_{s,\text{opt}}\|^2 \quad (6)$$

$$= \|\mathbf{w}_s^{(i)} - \mathbf{w}_{s,\text{opt}}\|^2 - 2\mu^{(i)} \mathbf{g}^{(i)\text{T}} \times (\mathbf{w}_s^{(i)} - \mathbf{w}_{s,\text{opt}}) + \mu^{(i)2} \|\mathbf{g}^{(i)}\|^2. \quad (7)$$

Note that $\mathbf{w}_{s,\text{opt}}$ is \mathbf{w}_{opt} with $(L+1)$ th element deleted. Inserting the relaxation step size rule

$$\mu^{(i)} = \gamma^{(i)} \frac{f^{(i)} - f_{\text{opt}}}{\|\mathbf{g}^{(i)}\|^2} \quad (8)$$

in (7), we obtain

$$\begin{aligned} \|\mathbf{w}^{(i+1)} - \mathbf{w}_{\text{opt}}\|^2 &= \|\mathbf{w}_s^{(i)} - \mathbf{w}_{s,\text{opt}}\|^2 - 2\gamma_i \frac{f^{(i)} - f_{\text{opt}}}{\|\mathbf{g}^{(i)}\|^2} \mathbf{g}^{(i)\text{T}} \\ &\quad \times (\mathbf{w}_s^{(i)} - \mathbf{w}_{s,\text{opt}}) + \gamma^{(i)2} \frac{(f^{(i)} - f_{\text{opt}})^2}{\|\mathbf{g}^{(i)}\|^2}. \end{aligned} \quad (9)$$

If we look at the inner product $\mathbf{g}^{(i)\text{T}}(\mathbf{w}_s^{(i)} - \mathbf{w}_{s,\text{opt}})$ in the second expression after equality sign in (9), we can write

$$\mathbf{g}^{(i)\text{T}} \mathbf{w}_s^{(i)} = f^{(i)} - \text{sign}(z_{l^{(i)}}^{(i)}) q_{l^{(i)}}. \quad (10)$$

Since \mathbf{g} is the channel output vector, we can write

$$\mathbf{g} = \text{sign}(z^{(i)}) \mathbf{C} \mathbf{x}, \quad (11)$$

where \mathbf{C} is obtained by deleting the $(L+1)$ th row of the $N_w \times (N_w + N_c - 1)$ channel convolution matrix

$$\mathbf{H} = \begin{bmatrix} c_0 & c_1 & \cdots & c_{N_c-1} & 0 & \cdots & 0 \\ 0 & c_0 & \cdots & c_{N_c-2} & c_{N_c-1} & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots \\ 0 & 0 & \cdots & c_0 & \cdots & c_{N_c-2} & c_{N_c-1} \end{bmatrix} \quad (12)$$

and \mathbf{x} is the channel input vector causing the maximum magnitude output. Since the maximum magnitude output is achieved when input samples take the extreme values from the constellation, x_i 's in (11) are $\pm Ms$ (since $\|z\|_\infty = \sum_{l=0}^{N_c-1} |c_{N_c-1-l}| M = \|c\|_1 M$ [3]).

We define

$$z_{\text{opt}}^{(i)} = \mathbf{Y}_{l^{(i)}}^{(i)} \mathbf{w}_{\text{opt}} \quad (13)$$

$$= \mathbf{x}^{(i)\text{T}} \mathbf{H}^{\text{T}} \mathbf{w}_{\text{opt}}, \quad (14)$$

where $\mathbf{x}^{(i)}$ is a vector formed by $\pm Ms$ and $\mathbf{H}^{\text{T}} \mathbf{w}_{\text{opt}}$, which corresponds to the convolution of the channel impulse response with the optimal equalizer coefficients, is approximately equal to a scaled delta function with magnitude f_{opt}/M . This approximation would be very reasonable as far as the convergence analysis is concerned especially for long equalizer lengths as the ISI will decrease with increasing equalizer lengths. When we consider large ISI levels (e.g. more than 20 dB) noise term that will enter in our convergence analysis due to this approximation will be negligible. Consequently, we can write the approximation as

$$z_{\text{opt}}^{(i)} \approx \text{sign}(z_{\text{opt}}^{(i)}) f_{\text{opt}} \quad (15)$$

from which we obtain

$$\Gamma_{l^{(i)}} \mathbf{w}_{s,\text{opt}} \approx \text{sign}(z_{\text{opt}}^{(i)}) f_{\text{opt}} - q_{l^{(i)}} \quad (16)$$

and therefore

$$\begin{aligned} \mathbf{g}^{(i)\text{T}} \mathbf{w}_{s,\text{opt}} &= \text{sign}(z^{(i)}) \Gamma_{l^{(i)}} \mathbf{w}_{s,\text{opt}} \\ &\approx \text{sign}(z^{(i)}) \text{sign}(z_{\text{opt}}^{(i)}) f_{\text{opt}} - \text{sign}(z^{(i)}) q_{l^{(i)}}. \end{aligned} \quad (17)$$

As a result, combining (10) with (17), the inner product in (9) can be approximated as

$$\mathbf{g}^{(i)\text{T}} (\mathbf{w}_s^{(i)} - \mathbf{w}_{s,\text{opt}}) \approx (f^{(i)} - \text{sign}(z^{(i)}) \text{sign}(z_{\text{opt}}^{(i)}) f_{\text{opt}}). \quad (18)$$

If $\text{sign}(z^{(i)}) = \text{sign}(z_{\text{opt}}^{(i)})$, which is true when $\mathbf{w}^{(i)}$ is at the vicinity of \mathbf{w}_{opt} , (9) can be rewritten as

$$\begin{aligned} \|\mathbf{w}^{(i+1)} - \mathbf{w}_{\text{opt}}\|^2 &\approx \|\mathbf{w}_s^{(i)} - \mathbf{w}_{s,\text{opt}}\|^2 - \gamma^{(i)}(2 - \gamma^{(i)}) \\ &\quad \times \frac{(f^{(i)} - f_{\text{opt}})^2}{\|\mathbf{g}^{(i)}\|^2} \\ &\approx \|\mathbf{w}_s^{(i)} - \mathbf{w}_{s,\text{opt}}\|^2 - \gamma^{(i)}(2 - \gamma^{(i)}) \\ &\quad \times \frac{|\mathbf{g}^{(i)\text{T}}(\mathbf{w}_s^{(i)} - \mathbf{w}_{s,\text{opt}})|^2}{\|\mathbf{g}^{(i)}\|^2} \end{aligned} \quad (19)$$

$$\begin{aligned} &\approx \|\mathbf{w}_s^{(i)} - \mathbf{w}_{s,\text{opt}}\|^2 \left(1 - \gamma^{(i)}(2 - \gamma^{(i)}) \right. \\ &\quad \left. \times \frac{|\mathbf{g}^{(i)\text{T}}(\mathbf{w}_s^{(i)} - \mathbf{w}_{s,\text{opt}})|^2}{\|\mathbf{g}^{(i)}\|^2 \|\mathbf{w}_s^{(i)} - \mathbf{w}_{s,\text{opt}}\|^2} \right). \end{aligned} \quad (20)$$

If Θ_i denotes the angle between the vectors $(\mathbf{w}_{\text{opt}} - \mathbf{w}^{(i)})$ and $\mathbf{g}^{(i)}$, then

$$|\cos(\Theta_i)| = \frac{|\mathbf{g}^{(i)\text{T}}(\mathbf{w}_s^{(i)} - \mathbf{w}_{s,\text{opt}})|}{\|\mathbf{g}^{(i)}\| \|\mathbf{w}_s^{(i)} - \mathbf{w}_{s,\text{opt}}\|}. \quad (21)$$

Therefore, with $\gamma^{(i)} = 1$, (20) simplifies to

$$\|\mathbf{w}^{(i+1)} - \mathbf{w}_{\text{opt}}\|^2 \approx \|\mathbf{w}^{(i)} - \mathbf{w}_{\text{opt}}\|^2 (1 - \cos^2(\Theta_i)) \quad (22)$$

$$\approx \|\mathbf{w}^{(i)} - \mathbf{w}_{\text{opt}}\|^2 \sin^2(\Theta_i). \quad (23)$$

The geometrical picture corresponding to the $\gamma^{(i)} = 1$ case is illustrated in Fig. 2, for a two dimensional \mathbf{w}_s . It can be seen from this figure that a right triangle is formed with the edges $\mathbf{w}_s^{(i+1)} - \mathbf{w}_{s,\text{opt}}$, $\mathbf{w}_s^{(i+1)} - \mathbf{w}_s^{(i)}$ and $\mathbf{w}_s^{(i)} - \mathbf{w}_{s,\text{opt}}$, where the right angle is between $\mathbf{w}_s^{(i+1)} - \mathbf{w}_s^{(i)}$ and $\mathbf{w}_s^{(i)} - \mathbf{w}_{s,\text{opt}}$. We should note that, for $\gamma^{(i)} \neq 1$, the resulting triangle would not be a right triangle.

In the case where $\text{sign}(z^{(i)}) \neq \text{sign}(z_{\text{opt}}^{(i)})$, then

$$\begin{aligned} \|\mathbf{w}^{(i+1)} - \mathbf{w}_{\text{opt}}\|^2 &\approx \|\mathbf{w}_s^{(i)} - \mathbf{w}_{s,\text{opt}}\|^2 \\ &\quad \times \left(1 - \gamma^{(i)}(2 - \gamma^{(i)}) \frac{|\mathbf{g}^{(i)\text{T}}(\mathbf{w}_s^{(i)} - \mathbf{w}_{s,\text{opt}})|^2}{\|\mathbf{g}^{(i)}\|^2 \|\mathbf{w}_s^{(i)} - \mathbf{w}_{s,\text{opt}}\|^2} \right) \\ &\quad - 4\gamma^{(i)}f_{\text{opt}} \frac{f^{(i)} - f_{\text{opt}}}{\|\mathbf{g}^{(i)}\|^2}, \end{aligned} \quad (24)$$

which corresponds to even faster convergence than (20), and for $\gamma^{(i)} = 1$,

$$\begin{aligned} \|\mathbf{w}^{(i+1)} - \mathbf{w}_{\text{opt}}\|^2 &\approx \|\mathbf{w}^{(i)} - \mathbf{w}_{\text{opt}}\|^2 \sin^2(\Theta_i) - 4f_{\text{opt}} \\ &\quad \times \frac{f^{(i)} - f_{\text{opt}}}{\|\mathbf{g}^{(i)}\|^2}. \end{aligned} \quad (25)$$

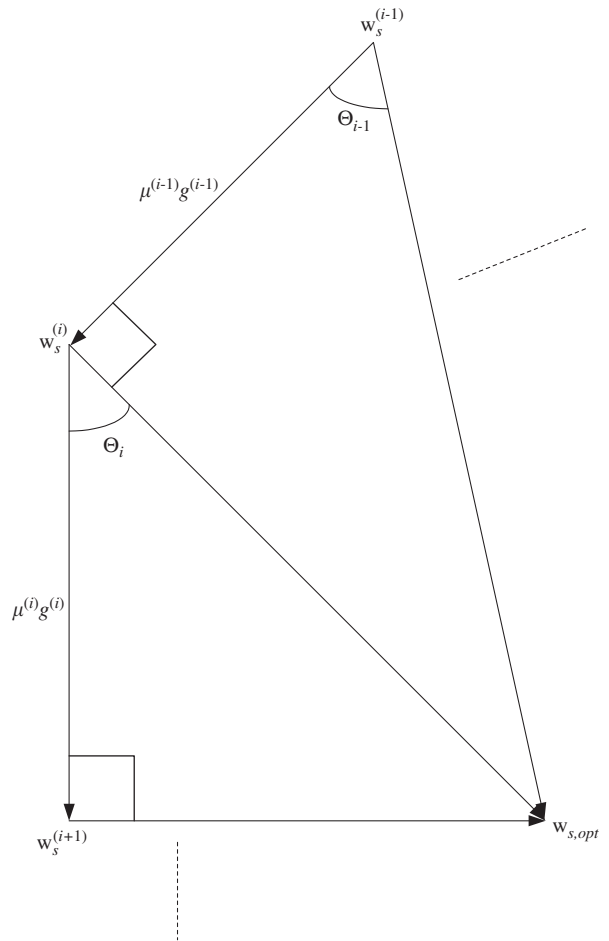


Fig. 2. The geometrical picture of updates for $\gamma_i = 1$.

4. Mean square deviation of the equalizer coefficients

In order to obtain a recursion expression for the mean square deviation of equalizer coefficient vector from the optimum value, we first take the expected value of both sides of (19) (where we implicitly assume that $\text{sign}(z^{(i)}) = \text{sign}(z_{\text{opt}}^{(i)})$ which holds true when $\mathbf{w}^{(i)}$ is in the vicinity of \mathbf{w}_{opt} and which gives a pessimistic rate of convergence):

$$\begin{aligned} E(\|\mathbf{w}^{(i+1)} - \mathbf{w}_{\text{opt}}\|^2) &\approx E(\|\mathbf{w}_s^{(i)} - \mathbf{w}_{s,\text{opt}}\|^2) \\ &\quad - \gamma^{(i)}(2 - \gamma^{(i)}) E\left(\frac{|\mathbf{g}^{(i)\text{T}}(\mathbf{w}_s^{(i)} - \mathbf{w}_{s,\text{opt}})|^2}{\|\mathbf{g}^{(i)}\|^2}\right). \end{aligned} \quad (26)$$

To get a more compact expression for the rightmost expectation, we assume that $\mathbf{g}^{(i)}$ and $\mathbf{w}_{s,\text{opt}} - \mathbf{w}_s^{(i)}$ are independent. This is a usual simplifying assumption

in the convergence analysis of various adaptive algorithms (e.g. LMS) (see for example, discussions in [8,9]), where the assumption is often justified by the slower change in $\mathbf{w}_{s,\text{opt}} - \mathbf{w}_s^{(i)}$ in comparison to the change in $\mathbf{g}^{(i)}$. Although this assumption may not be strictly correct, the simulation results in Section 5 show that the analysis based on this assumption provides a close characterization of the convergence behavior. Under this assumption we can write

$$E\left(\frac{|\mathbf{g}^{(i)\text{T}}(\mathbf{w}_s^{(i)} - \mathbf{w}_{s,\text{opt}})|^2}{\|\mathbf{g}^{(i)}\|^2}\right) \approx E((\mathbf{w}_s^{(i)} - \mathbf{w}_{s,\text{opt}})^{\text{T}} E\left(\frac{\mathbf{g}^{(i)}\mathbf{g}^{(i)\text{T}}}{\mathbf{g}^{(i)\text{T}}\mathbf{g}^{(i)}}\right) (\mathbf{w}_s^{(i)} - \mathbf{w}_{s,\text{opt}})) \quad (27)$$

from which we obtain

$$\begin{aligned} \lambda_{\min}\left(E\left(\frac{\mathbf{g}^{(i)}\mathbf{g}^{(i)\text{T}}}{\|\mathbf{g}^{(i)}\|^2}\right)\right) E(\|\mathbf{w}^{(i)} - \mathbf{w}_{\text{opt}}\|^2) \\ \leq E\left(\frac{|\mathbf{g}^{(i)\text{T}}(\mathbf{w}^{(i)} - \mathbf{w}_{\text{opt}})|^2}{\|\mathbf{g}^{(i)}\|^2}\right) \\ \leq \lambda_{\max}\left(E\left(\frac{\mathbf{g}^{(i)}\mathbf{g}^{(i)\text{T}}}{\|\mathbf{g}^{(i)}\|^2}\right)\right) E(\|\mathbf{w}^{(i)} - \mathbf{w}_{\text{opt}}\|^2). \end{aligned} \quad (28)$$

We make the following approximation to obtain a more explicit expression for the convergence rate bounds:

$$E\left(\frac{\mathbf{g}^{(i)}\mathbf{g}^{(i)\text{T}}}{\|\mathbf{g}^{(i)}\|^2}\right) \approx \frac{E(\mathbf{g}^{(i)}\mathbf{g}^{(i)\text{T}})}{E(\|\mathbf{g}^{(i)}\|^2)} = \frac{\mathbf{R}_g}{\sigma_g^2}. \quad (29)$$

Hence, we can rewrite the range in (28) as

$$\frac{\lambda_{\min}(\mathbf{R}_g)}{\sigma_g^2} \leq E\left(\frac{|\mathbf{g}^{(i)\text{T}}(\mathbf{w}^{(i)} - \mathbf{w}_{\text{opt}})|^2}{\|\mathbf{g}^{(i)}\|^2}\right) \leq \frac{\lambda_{\max}(\mathbf{R}_g)}{\sigma_g^2}. \quad (30)$$

As stated previously, \mathbf{g} is generated by \mathbf{x} vector whose elements are $\pm Ms$. Furthermore, we assume that the elements of \mathbf{x} are uncorrelated with each other. Based on this data model the covariance matrix \mathbf{R}_g can be written as

$$\mathbf{R}_g = M^2 \mathbf{C} \mathbf{C}^{\text{T}} \quad (31)$$

and

$$\sigma_g^2 = M^2 \text{Tr}\{\mathbf{C} \mathbf{C}^{\text{T}}\}. \quad (32)$$

As a result, we can write the recursion for the mean square deviation of the equalizer coefficients as

$$E(\|\mathbf{w}^{(i+1)} - \mathbf{w}_{\text{opt}}\|^2) = E(\|\mathbf{w}^{(i)} - \mathbf{w}_{\text{opt}}\|^2) \times (1 - \cos^2(\Theta_i)), \quad (33)$$

where

$$\frac{\lambda_{\min}(\mathbf{C} \mathbf{C}^{\text{T}})}{\text{Tr}\{\mathbf{C} \mathbf{C}^{\text{T}}\}} \leq \cos^2(\Theta_i) \leq \frac{\lambda_{\max}(\mathbf{C} \mathbf{C}^{\text{T}})}{\text{Tr}\{\mathbf{C} \mathbf{C}^{\text{T}}\}}. \quad (34)$$

Alternatively, in terms of the singular values of the channel convolution matrix, we can write

$$\begin{aligned} \frac{\sigma_{\min}(\mathbf{C})}{\sqrt{\sigma_{\min}^2(\mathbf{C}) + \dots + \sigma_{\max}^2(\mathbf{C})}} \\ \leq |\cos(\Theta_i)| \\ \leq \frac{\sigma_{\max}(\mathbf{C})}{\sqrt{\sigma_{\min}^2(\mathbf{C}) + \dots + \sigma_{\max}^2(\mathbf{C})}}. \end{aligned} \quad (35)$$

5. Examples

In this section, we will test the validity of the bounds obtained by our mean square deviation analysis in Section 4. For that purpose, we simulate the SGBA algorithm using the relaxation step size rule for some sample channels. In these simulations, we record the distance to the optimal point at each iteration. Upper bound and lower bound curves are obtained by using the best case and worst case geometric convergence rates suggested by (35).

For the first example, we consider “the good quality phone channel” in [10] whose impulse response is shown in Fig. 3. The distance-to-optimal-point convergence curve for this channel is shown in Fig. 4, where the upper and lower bounds obtained in the previous section are also shown. It is clear from this figure that the convergence curve closely follows the lower (fast convergence) curve at the initial segment of the convergence process, and then it becomes parallel to the upper (slow convergence) curve. Therefore, we can conclude that the convergence behavior of the SGBA for this sample channel fits well to our analysis results.

As the second example, we consider an autoregressive channel with transfer function

$$H(z) = \frac{1}{1 + 0.3z^{-1} - 0.28z^{-2}}. \quad (36)$$

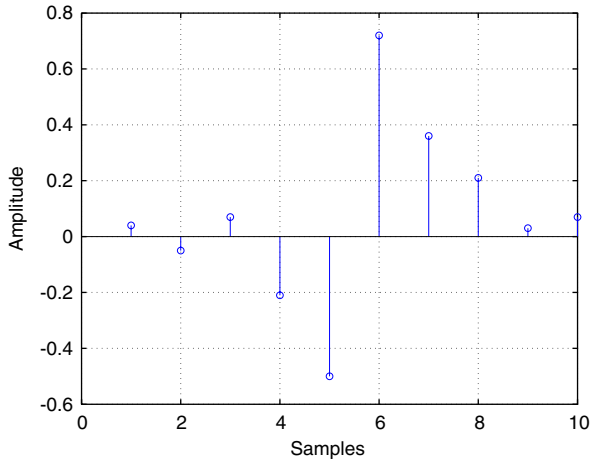


Fig. 3. Impulse response of test channel 1.

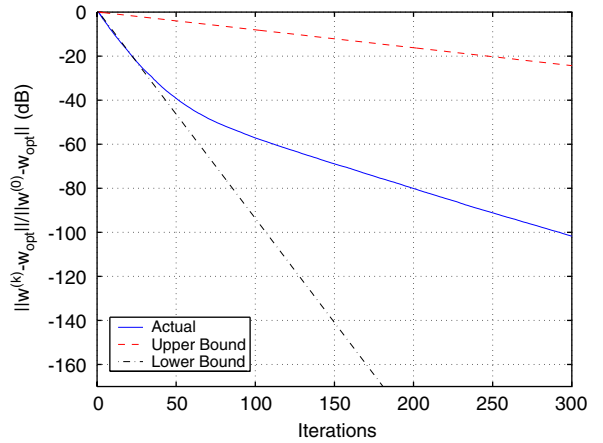


Fig. 5. Convergence curve for the test channel 2.

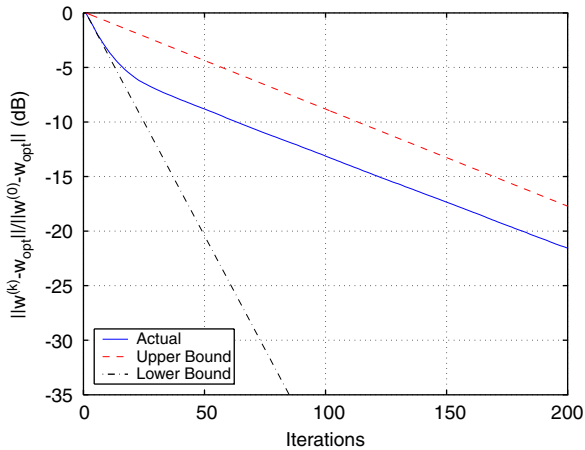


Fig. 4. Convergence curve for the test channel 1.

The convergence curve for this example is shown in Fig. 5. Similar to the previous example, the convergence curve first follows the fast convergence curve and then the convergence rate decreases.

Based on both examples, and various simulation trials for some example channels, we can conclude that upper and lower bound convergence curves obtained by (35) provide reasonable bounds for the convergence behavior of the SGBA.

6. Conclusion

In this article, we provided a convergence analysis for the SGBA algorithm. We first obtained the

geometric picture corresponding to the updates, where the monotonic decrease in distance to the optimal equalizer can be easily visualized. The scale of reduction in distance is dependent on the angle between the current error and the corresponding subgradient vectors.

We also analyzed the mean square convergence behavior for the distance to the optimal point by making some reasonable approximations. Based on this analysis, we could show that the convergence curve is upper and lower bounded by two geometric-series curves whose geometric factors are dependent on the singular values of the channel mapping matrix corresponding to the variable taps.

The simulation results validate both the smooth-monotonic convergence behavior of the SGBA algorithm and the applicability of the geometric-series bounds obtained by our analysis.

References

- [1] S. Lambotharan, J. Chambers, C.R. Johnson, Attraction of saddles and slow convergence in CMA adaptation, *Signal Processing* 59 (1997) 335–340.
- [2] Z. Ding, R. Kennedy, B. Anderson, C.R. Johnson, Ill-convergence of godard blind equalizers in data communication systems, *IEEE Trans. Comm.* 39 (1991) 1313–1327.
- [3] S. Vembu, S. Verdu, R. Kennedy, W. Sethares, Convex cost functions in blind equalization, *IEEE Trans. Signal Process.* 42 (1994) 1952–1960.
- [4] Z. Ding, Z. Luo, A fast linear programming algorithm for blind equalization, *IEEE Trans. Comm.* 48 (2000) 1432–1436.
- [5] Z.-Q. Luo, M. Meng, K.M. Wong, J.-K. Zhang, A fractionally spaced blind equalizer based on linear programming, *IEEE Trans. Signal Process.* 50 (2002) 1650–1660.

- [6] A.T. Erdogan, C. Kizilkale, Fast and low complexity blind equalization via subgradient projections, *IEEE Trans. Signal Process.* 53 (2005) 2513–2524.
- [7] C. Kizilkale, A.T. Erdogan, A fast blind equalization method based on subgradient projections, *Proceedings of IEEE ICASSP 2004, Montreal, Canada*, vol. 4, pp. 873–876.
- [8] S. Haykin, *Adaptive Filter Theory*, Prentice-Hall, Englewood Cliffs, NJ, 2002.
- [9] A.H. Sayed, *Fundamentals of Adaptive Filtering*, Wiley, New York, 2003.
- [10] J.G. Proakis, *Digital Communications*, McGraw-Hill, New York, 2001.