ON PROPERTIES OF ELECTRONIC EQUALIZERS FOR OPTICAL COMMUNICATIONS SYSTEMS

Wei Wang, Tülay Adalı, Wenze Xi, and Curtis R. Menyuk

Department of Computer Science and Electrical Engineering
University of Maryland Baltimore County, Baltimore, MD, 21250
Tel: (410) 455-1631 e-mail: wangw1@csee.umbc.edu

ABSTRACT

We study the properties of electronic equalizers for optical communications systems and derive expressions that explain their behavior. We use simulations to show the agreement with the theory developed.

1. INTRODUCTION

Electronic equalization is becoming increasingly important for mitigation of physical impairments in optical communication systems [1]. In particular, it has been shown that it is effective in reducing the penalty due to polarization mode dispersion (PMD), an important limiting factor for transmission rates and distances in installed terrestrial fiber systems. Experimental results at 10 Gbit/s use SiGe and GaAs technology and report successful results for PMD mitigation [2] (also see [1] and the references cited therein).

These equalizers are primarily based on the mean square error (MSE) criterion and minimize the square of the prediction error (difference between the equalizer output and the undistorted, *i.e.*, desired, signal). Solution to the MSE minimization problem is given by the Wiener-Hopf equations and can be adaptively estimated using gradient descent minimization. The least mean squares (LMS) algorithm uses an instantaneous estimate of the statistics required for the gradient updates of the filter coefficients, and is a simple and effective solution that has been successfully used in most adaptive filtering applications to date. It has also been applied to equalization solutions for the optical domain along with gradient descent optimization and tuning of the filter coefficients so that they satisfy the Wiener-Hopf equations.

Certain properties of the optical domain however are different than transmission media such as wireless and wireline (telephone) networks where the MSE equalizers have been widely used. In optical systems, bipolar signal transmission formats are seldom used, hence the transmitted signal is non-zero mean and the use of direct detection at the receiver introduces nonzero mean and signal-dependent noise into the received signal, the input of the electronic equalizer.

Autocorrelation matrix of the received signal describes the behavior of techniques such as LMS that rely on gradient optimization of the MSE cost. In this paper, we derive expressions describing dependence of eigenvalues of the correlation matrix on system parameters and show the deviation from the minimum MSE that can be achieved in the presence of nonzero mean noise and signal. We assume transversal (feed-forward) type filter characteristics, and use discrete notation in the analysis. The conclusions however extend to analog implementations of the MSE equalizer. We discuss the significance of these results and use simulations to demonstrate the agreement with the theory.

2. INPUT CORRELATION MATRIX AND ITS EIGENANALYSIS

The correlation matrix, and in particular, its eigenvalues determine the performance of filters based on the MSE criterion. A high eigenvalue disparity measured by the ratio of the maximum eigenvalue to the minimum one implies numerical stability problems as the Wiener-Hopf solution requires the computation of the inverse of this matrix. In gradient optimization based computation of this solution such as by using LMS or other methods such as gradient minimization by eye-monitoring [1], the convergence of the algorithm depends on the absolute value of the eigenvalues, small eigenvalues resulting in "slow modes" for the adaptation.

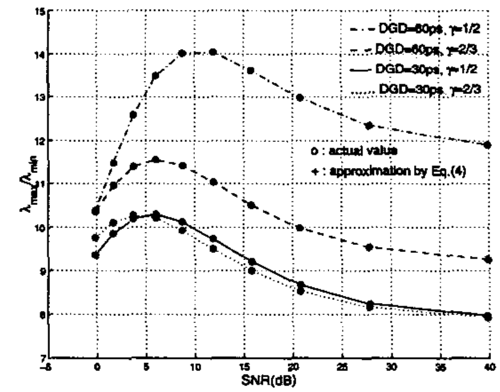

Fig. 1. Approximation of the eigenvalue spread of **R** by equation (4). A first order PMD channel is simulated with the given γ (power split ratio of the principal states of polarization) and the differential group delay (DGD) values and with the following parameters: Gaussian return-to-zero pulses; Peak power=1mW, Bit duration=100ps, FWHM=50ps.

To observe the change in the correlation matrix $\mathbf{R}$ with different noise and input mean values, let $v(n)$ denote the noiseless input to the adaptive filter, and $u(n) = v(n) + \omega(n)$ the noisy input with $\omega(n) \sim N(\mu_\omega, \sigma^2)$ and $E[v(n)] = \mu_v$. The autocorrelation matrix of $u(n)$, $\mathbf{R} = E[\mathbf{u}(n)\mathbf{u}^H(n)]$, is written as:

$$\mathbf{R} = \mathbf{R}_0 + \alpha\mathbf{I} + \beta\mathbf{1}\mathbf{1}^T \qquad (1)$$

where $\mathbf{R}_0 = E[\mathbf{v}(n)\mathbf{v}^H(n)]$, $\alpha = \sigma^2$, $\beta = \mu_\omega^2 + 2\mu_v\mu_\omega$, $\mathbf{1}$ is an $M \times 1$ dimensional vector with all 1s, and the boldface quantities denote the vectors that contain the last M observations, e.g., $\mathbf{v}(n) = [v(n), v(n-1), \ldots, v(n-M+1)]^H$. The dimension of the transversal filter is assumed to be M.

The eigenvalues of $\mathbf{R}$ can be obtained by solving a *secular equation* [3], and for an input and noise with nonzero mean we can obtain reliable approximations for the eigenvalues of $\mathbf{R}$. For example, the approximate maximum eigenvalue is given by:

$$\hat{\lambda}_{\max} = \lambda^0_{\max} + \alpha + \frac{\beta q_1^2}{1 - \beta\sum_{i=2}^{M} \frac{q_i^2}{\lambda^0_{\max} - \lambda^0_i}} > \lambda_{\max} \approx \lambda^0_{\max} + \alpha + \beta q_1^2 \qquad (2)$$

where λ_i^0 denotes the eigenvalues of $\mathbf{R_0}$, such that $\lambda_1^0 > \lambda_2^0 > \ldots > \lambda_M^0$, $\mathbf{q} \equiv \mathbf{Q}^H\mathbf{1} = [q_1, q_2, \ldots, q_M]^T$, $\mathbf{Q}$ is an $M \times M$ unitary matrix with columns as the orthonormal set of eigenvectors, and $\lambda_{\max}$ is the actual maximum eigenvalue of $\mathbf{R}$. The minimum eigenvalue of $\mathbf{R}$ can also be computed as

$$\hat{\lambda}_{\min} = \lambda_{\min}^0 + \alpha + \frac{\beta q_n^2}{1 + \beta \sum_{i=1}^{M-1} \frac{q_i^2}{\lambda_i - \lambda_{min}}} \approx \lambda_{\min}^0 + \alpha \qquad (3)$$

Thus we obtain the spread of $\mathbf{R}$ as:

$$\chi(\mathbf{R}) = \lambda_{\max}/\lambda_{\min} \approx \frac{\lambda_{\max}^0 + \alpha + \beta q_1^2}{\lambda_{\min}^0 + \alpha} \qquad (4)$$

Fig. 1 shows the plot of the eigenvalue spread of $\mathbf{R}$ as a function of the signal to noise ratio (SNR) defined as the ratio of the signal power to the noise power. As observed in the figure, the given approximations are quite close to the actual values and the spread increases with increasing noise levels and when the input and noise are nonzero mean. This increases the sensitivity of the solution and slows down the convergence of the gradient descent based adaptive learning algorithm. Note that as the noise in the system increases there is a point after which the spread starts to decrease as then, noise starts to dominate in the statistics improving the conditioning of the matrix. Also, the increase in the maximum eigenvalue with nonzero mean noise characteristics as observed in equation (2) implies increased sensitivity to the choice of the step size in an optical system.

3. MINIMUM MSE AND GRADIENT OPTIMIZATION PERFORMANCE

Filter coefficients minimizing the MSE, $\mathbf{w}_{\text{opt}}$, are given by the solution of Wiener-Hopf equations: $\mathbf{R}\mathbf{w}_{\text{opt}} = \mathbf{p}$, where $\mathbf{p}$ is the cross-correlation vector between the inputs of the equalizer $u(n)$s and the desired signal $d(n)$, and $\mathbf{p} = \mathbf{p}_0 + \mu_d\mu_\omega\mathbf{1}$, where $\mathbf{p}_0 = E[\mathbf{v}(n)d^*(n)]$ with * denoting the complex conjugate. The minimum MSE, $\mathbf{J}_{\min}$, is achieved when the filter coefficients are given by $\mathbf{w}_{opt}$ (either computed by tuning the coefficients for weighting the delayed input samples or by an adaptive gradient-descent type procedure) is given by

$$\mathbf{J}_{\min} = \sigma_d^2 + \mu_d^2 - \mathbf{p}^H\mathbf{R}^{-1}\mathbf{p} \qquad (5)$$

where μ_d and σ_d are the mean and variance of $d(n)$, respectively. To obtain an expression for the increase in $\mathbf{J}_{\min}$ with nonzero mean and input, use the Sherman-Morrison formula [3], to write the inverse of $\mathbf{R}$:

$$\mathbf{R}^{-1} = \mathbf{Q}\mathbf{\Lambda}^{-1}\mathbf{Q}^H - \frac{\beta\mathbf{Q}\mathbf{\Lambda}^{-1}\mathbf{q}\mathbf{q}^H\mathbf{\Lambda}^{-1}\mathbf{Q}^H}{1 + \beta\sum\frac{q_i^2}{\lambda_i}} \qquad (6)$$

$$\text{where } \mathbf{\Lambda}^{-1} = \begin{pmatrix} \ddots & 0 & 0 \\ 0 & \frac{1}{\lambda_i^0 + \alpha} & 0 \\ 0 & 0 & \ddots \end{pmatrix} \qquad (7)$$

The change in the minimum MSE with nonzero mean characteristics is then written as:

$$\Delta\mathbf{J}_{\min} \equiv \mathbf{J}_{\min} - \mathbf{J}_{\min}^0 = \mathbf{p}_0^H\mathbf{R}_0^{-1}\mathbf{p}_0 - \mathbf{p}^H\mathbf{R}^{-1}\mathbf{p} \qquad (8)$$

where $\mathbf{J}_{\min}^0$ is the minimum MSE without noise. Fig. 2 shows the increase in the misadjustment, $\Delta\mathbf{J}_{\min}$, as a function of the SNR. The misadjustment decreasing with SNR, increasing DGD values, and as the splitting factor γ for a first order PMD channel approaches 0.5. The upper curve also shows the comparison of the expressions derived with simulated results. The close agreement observed for this case holds in general, and when the multiplicative noise that is ignored in the current analysis is present in the system as well. These results are not included due to space limitations.

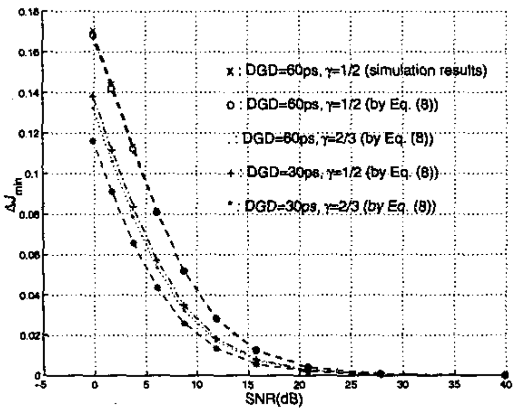

Fig. 2. Change in minimum MSE $\Delta\mathbf{J}_{min}$ with different noise levels for the same parameters as in Fig. 1.

In gradient descent optimization, filter coefficients are updated by moving in the opposite direction of the gradient, maximum change in MSE, and the step size μ is an important parameter controlling the rate of adaption, the final misadjustment, and the stability of the updates.

Substituting equation (2) into the well known convergence criterion for gradient optimization of MSE [4]: $0 < \mu < 2/\lambda_{\max}$, we obtain the following stability condition:

$$0 < \mu < \frac{2}{\lambda_{\max}^0 + \sigma^2 + (\mu_\omega^2 + 2\mu_v\mu_\omega)q_1^2} \qquad (9)$$

Hence, the presence of nonzero noise (μ_ω) and input (μ_v) decreases the upper bound for the step size, increasing the sensitivity of the algorithm to its choice. A choice of $\mu = 2/(\lambda_{\max} + \lambda_{\min})$ provides fast convergence for the updates.

4. CONCLUSIONS

We derive expressions to explain the properties of the most widely employed electronic equalizers, those based on MSE minimization, when used in optical systems. We show that the statistics of input and noise present in the optical system introduce bias into the estimates and might significantly slow down the convergence of gradient descent-based minimization, such as LMS, the most common and practical way for adaptive computation of the optimum minimum MSE filter coefficients. The nonzero mean of the input leads to increased eigenvalue spread of the input correlation matrix slowing down the convergence. In this case, subtracting the mean of the input prior to equalization will improve the convergence of the gradient-descent based adaptation algorithm [5].

5. REFERENCES

[1] H. Bülow, "Electronic equalization of transmission impairments", in *OFC'2002 Tech. Digest*, Anaheim, CA, Mar. 2002, TuF.

[2] H. Bülow, *et al.*,"PMD mitigation at 10 Gbit/s using linear and nonlinear integrated electronic equalizer circuits," *Electron. Letts.*, vol. 36, no. 2, pp. 163–164, Jan. 2000.

[3] J. W. Demmel, Applied Numerical Linear Algebra, SIAM 1997.

[4] S. Haykin, Adaptive Filter Theory, Prentice Hall, 2002.

[5] J.-N. Lin, R. Unbehauen,"Adaptive filtering incorporating a local mean estimation substructure", *IEE Proceedings G*, vol. 140, no. 1, pp. 16–22, Feb. 1993.