FAST ALGORITHM FOR FINITE-LENGTH MMSE EQUALIZERS WITH APPLICATION TO DISCRETE MULTITONE SYSTEMS

Navid Lashkarian

ECE Department
Oregon State University
Email: lashkadd@ece.orst.edu

Sayfe Kiaei

Motorola, Inc. SPS.
Austin, TX
Email: esk013@email.sps.mot.com

ABSTRACT

This paper presents a *new*, fast algorithm for finite-length minimum mean square error (MMSE) equalizers. The research exploits asymptotic equivalence of Toeplitz and circulant matrices to estimate Hessian matrix of a quadratic form. Research shows that the Hessian matrix exhibits a specific structure. As a result, when combined with the Rayleigh minimization algorithm, it provides an efficient method to obtain the global minimum of constrained optimization problem. A salient feature of this algorithm is that extreme eigenvector of the Hessian matrix can be obtaind without direct computation of the matrix. In comparison to the previous methods, the algorithm is more computationally efficient and highly parallelizable, which makes the algorithm more attractive for real time applications. The algorithm is applied for equalization of discrete multitone (DMT) systems for asynchronous digital subscriber line (ADSL) applications.

1. INTRODUCTION

In design of adaptive filters for signal processing applications, various optimality criteria can be used to obtain the optimum setting of the adaptive filter. However, MMSE is considered to be the most tractable technique which guarantees existence and uniqueness of global optimum solution. The problem of finite-length MMSE filtering has already been investigated in many literatures [1],[2]. In [2], author applies the notion of MMSE filtering for system identification problems. Partial equalization of spectrally shaped channels is another fruitful application of MMSE filtering in communication and signal processing. Specifically, given a highly disperive channel of length v, the objective is to design a finite-length time domain equalizer (TEQ) to force the effective channel into a much shorter filter known as target impulse response (TIR). In general, the optimum solution to this problem is obtained from computing the global minimum of a quadratic function. Due to the inherent potential of quadratic forms to converge to the trivial solution, an energy boosting constraint is applied to the problem. Among the feasible constraint sets, unit energy constraint (UEC) and unit tap constraint (UTC) have found more applications in communication systems. In principle, decision feedback equalization (DFE) can be categorized as a special class of MMSE equalizers under UTC. A fast algorithm for MMSE equalizers has already been proposed in [4]. However, the study conducted in [4] provided the optimum solution subject to UTC. In this paper, a new fast iterative algorithm for computing optimum setting of MMSE-UEC equalizers is presented. Also note that equalization under UEC provides better SNR in comparison to UTC [1]. The method makes use of asymptotic equivalence of circulant and Toeplitz matrices to obtain a closed form expression for the Hessian matrix. Additionally, we show that any quadratic form can be computed efficiently using the discrete Fourier transform (DFT) operation. When combined with the Rayleigh minimization algorithm, it provides a fast algorithm for computing coefficients of TIR and TEQ . The algorithm provides the solution after $N_b + 1$ iterations and requires $O(N_f log_2(N_f))$ operations/iteration where N_f and $N_b + 1$ are the length of TEQ and TIR, respectively. The rest of this paper is organized as follows. In section 2 an overview of MMSE approach is presented. In section 3 few proporties of Hessian matrix are derived. Based on these derivations, a new iterative algorithm for MMSE-UEC is proposed. The complexity of algorithm is compared against the standard matrix inversion method. Finally, in section 4 the algorithm is applied to impulse response shortening of DMT systems.

2. MMSE EQUALIZATION

This section presents an overview of the MMSE equalization problem.[1] Block diagram of the equalizer studied in this paper is depicted in figure 1. The channel response is modeled as a discrete time FIR filter, expressed by $\mathbf{h} = \{\mathbf{h}[0], \mathbf{h}[1], \cdots, \mathbf{h}[v]\}$ where v is the channel spread. The channel response represents the combined effect of the transmit and receive filters as well as the channel impulse response. Input is an independent identically distributed random sequence with power of σ_x^2. In MMSE approach, equalizer taps are set such that the residual error between output of TIR and TEQ filters is minimized in the mean square sense. MMSE equalization can be viewed as a quadratic optimization problem in which the optimum settings for the TIR and TEQ filters are obtained from the following equations

$$\mathbf{b}_{opt} = \arg \min_{\mathbf{b}} \ \mathbf{b}^* \mathbf{R}_\triangle \mathbf{b} \tag{1}$$

$$\mathbf{w} = \sigma_x^2 \mathbf{R_{yy}}^{-1} \mathbf{H}_\triangle \mathbf{b}_{opt} \tag{2}$$

where $\mathbf{R}_\triangle$ is the Hessian matrix given in

$$\mathbf{R}_\triangle \stackrel{\text{def}}{=} \sigma_x^2 \mathbf{I}_{\mathbf{N_b}+1} - \sigma_\mathbf{x}^4 \mathbf{H}_\triangle^* \mathbf{R}_{\mathbf{yy}}^{-1} \mathbf{H}_\triangle \tag{3}$$

and

$$\mathbf{R}_{yy} \quad = \quad \sigma_x^2 \mathbf{H}\mathbf{H}^* + \mathbf{R}_{nn}$$

[1] Throughout the paper, symbols $\odot$, $*$, $\Re$ and $\tilde{\cdot}$, represent element by element vector multiplication, linear convolution, real and Fourier transform operations respectively. Also matrices and vectors are represented by uppercase and lowercase bold characters, respectively.

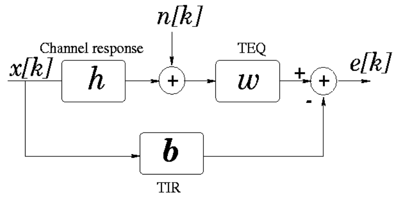

Figure 1: Block Diagram of MMSE Equalizer

$$\mathbf{H}_{\boldsymbol{\Delta}} = \mathbf{HS}^*$$
$$\mathbf{H} = \begin{bmatrix} \mathbf{h}[0] & \mathbf{h}[1] & \cdots & \mathbf{h}[v] & \cdots & 0 \\ 0 & \mathbf{h}[0] & \mathbf{h}[1] & \cdots & \mathbf{h}[v] & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & \mathbf{h}[0] & \mathbf{h}[1] & \cdots & \mathbf{h}[v] \end{bmatrix}$$
$$\mathbf{S} = \begin{bmatrix} \mathbf{0}_{N_b+1,\Delta} & \mathbf{I}_{N_b+1,N_b+1} & \mathbf{0}_{N_b+1,\kappa} \end{bmatrix}$$

In the above equations, Δ is the decision delay involved with the TIR and $\kappa \stackrel{\text{def}}{=} N_f + v - N_b - 1$. Also matrices $\mathbf{I}$,$\mathbf{0}$ and $\mathbf{R}_{nn}$ represent the identity,zero and noise autocorrelation matrices, respectively. Recall that in the presence of UEC, the optimization problem given in (1) can be viewed as minimization of Rayleigh quotient of matrix $\mathbf{H}_{\boldsymbol{\Delta}}$. With this idea in mind, we apply the iterative algorithm proposed in [5] to obtain extreme eigenvalue (vector) of the Hessian matrix. Basically, this algorithm applies the conjugate-gradient method to find the minimum of second order approximation of Rayleigh quotient. Rayleigh minimization algorithm is widely used in subsapce tracking problems [7]. This method, although iterative in nature, is guaranteed to converge after at most $N_b + 1$ iterations. Robustness to matrix condition number is another noticeable feature of this algorithm [6]. Although the Rayleigh minimization algorithm does not exploit matrix inversion , it requires frequent computation of quadratic forms. In the next section we describe an efficient method for computing the quadratic terms, which eventually leads us to a fast algorithm for computing coefficients of MMSE-UEC equalizers.

3. ITERATIVE ALGORITHM

The first step in reducing the complexity of the algorithm is to provide an efficient method for computing the inverse of the autocorrelation matrix. In doing so, we approximate the Toeplitz matrix

$$\mathbf{R}_{yy} = Tpltz(r_{yy}[0], r_{yy}[1], \quad \cdots \quad, r_{yy}[N_f - 1])$$

with its asymptotic equivalent, known as circulant matrix

$$\mathbf{C}_{yy} = Tpltz(r_{yy}[0], r_{yy}[1], \cdots, r_{yy}[N_f/2], \\ r_{yy}[N_f/2 - 1], \cdots, r_{yy}[1])$$

Note that the argument of the $Tpltz$ operator is the first row of the Toeplitz matrix. A distinct advantage of circulant matrix is that it can be decomposed into the product of Fourier matrices and a diagonal matrix as given by

$$\mathbf{C}_{yy} = \mathbf{U}_{N_f} \ \Psi \ \mathbf{U}^*_{N_f}$$

where

$$\Psi = diag(\psi[0] \quad \psi[1] \quad \cdots \quad \psi[N_f - 1])$$
$$\mathbf{U}_{N_f}[k,l] = \frac{1}{\sqrt{N_f}} e^{-j\frac{2\pi kl}{N_f}} \quad k,l = 0,1,...N_f - 1$$
$$\psi[k] = \sum_{m=0}^{N_f-1} \mathbf{c}[m] e^{-j\frac{2\pi km}{N_f}} \quad k = 0,1,...N_f - 1$$

and $\mathbf{c}[j]$ is the j'th element of the first row of matrix $\mathbf{C}_{yy}$. Using orthogonal propoerties of Fourier matrices, we can estimate the Hessian matrix as

$$\mathbf{R}_\Delta = {\sigma_x}^2 \mathbf{I}_{N_b+1} - {\sigma_x}^4 \mathbf{H}^*_\Delta \mathbf{U}^*_{N_f} \ \Psi^{-1} \ \mathbf{U}_{N_f} \ \mathbf{H}_\Delta \tag{4}$$

This closed form expression appears to attain fruitful propoerties as we will illustrate shortly.

Property 1:

Given an arbitrary pair of vectors $\mathbf{p}$ and $\mathbf{q}$ of length $N_b + 1$, the quadratic term $\mathbf{p}^* \mathbf{R}_\Delta \mathbf{q}$ can be expressed as

$$\mathbf{p}^* \mathbf{R}_\Delta \mathbf{q} = \mathbf{p}^* ({\sigma_x}^2 \mathbf{I}_{N_b+1} - {\sigma_x}^4 \mathbf{H}^*_\Delta \mathbf{U}^* \ \Psi^{-1} \ \mathbf{U} \ \mathbf{H}_\Delta)\mathbf{q} \tag{5}$$

Define a dummy vector $\mathbf{c} \stackrel{\text{def}}{=} {\sigma_x}^2 \mathbf{H}_\Delta \mathbf{q}$. Due to the circular property of matrix $\mathbf{H}_\Delta$ this vector can be written as linear convolution of two vectors as expressed by

$$\mathbf{c}[n] = {\sigma_x}^2 \sum_{l=0}^{N_f-1} \mathbf{h}[\Delta - n + l]\, \mathbf{q}[l] = {\sigma_x}^2 \mathbf{g}[n] * \mathbf{q}[n]$$

where $\mathbf{g}[n] \stackrel{\text{def}}{=} \mathbf{h}[-n + \Delta]$. The term ${\sigma_x}^2 \mathbf{U} \mathbf{H}_\Delta \mathbf{q} = \tilde{\mathbf{c}}$ is simply the Fourier transform of vector $\mathbf{c}$ and can be computed efficiently as

$$\sigma_x^2 \mathbf{U} \mathbf{H}_\Delta \mathbf{q} = \tilde{\mathbf{c}} = \sigma_x^2 \tilde{\mathbf{g}} \odot \tilde{\mathbf{q}} \tag{6}$$

Note that in the above expression, we have assumed that the length of TEQ exceeds that of TIR filter ($N_f > N_b + 1$). In applications in which this constraint can not be tolerated, the long TEQ filter can be well approximated by a pole-zero filter with fewer coefficients [3]. Applying equation (6) and Parseval's equality to equation (5) results in a closed form expression for the quadratic term $\mathbf{p}^* \mathbf{R}_\Delta \mathbf{q}$ as given by

$$\mathbf{p}^* \mathbf{R}_\Delta \mathbf{q} = \sum_{k=0}^{N_f-1} \tilde{\mathbf{z}}[k]\, \tilde{\mathbf{p}}^*[k]\, \tilde{\mathbf{q}}[k] \tag{7}$$

where the new vector $\tilde{\mathbf{z}}$ is defined as

$$\tilde{\mathbf{z}}[k] \stackrel{\text{def}}{=} {\sigma_x}^2 - {\sigma_x}^4 \frac{\tilde{\mathbf{g}}[k] \tilde{\mathbf{g}}^*[k]}{\psi[k]} \tag{8}$$

This closed form expression given in (7) suggests performing the Rayleigh minimization algorithm in the frequency domain. In doing so, we need to represent the Fourier transform of vector $\mathbf{R}_\Delta \mathbf{q}$ as a function of vector $\tilde{\mathbf{q}}$. Wishing to avoid performing the above operation in the time domain, we propose an efficient method which performs the above operation using DFT.

Property 2:

For a vector $s \stackrel{\text{def}}{=} \mathbf{R}_\Delta \mathbf{q}$, the i'th element can be represented as

$$s[i] = \mathbf{e}_i \mathbf{R}_\Delta \mathbf{q} = \frac{1}{\sqrt{N_f}} \sum_{k=0}^{N_f-1} \tilde{\mathbf{z}}[k]\, \tilde{\mathbf{q}}[k]\, e^{j\frac{2\pi ki}{N_f}} \tag{9}$$

where $\mathbf{e}_i$ is the i'th unit vector of length $N_b + 1$. In deriving the above equation, we have used the closed form expression given in (7). Equation (9) appears to be the i'th element of IDFT of vector $\tilde{\mathbf{z}} \odot \tilde{\mathbf{q}}$. Hence, the vector $\mathbf{s}$ can be obtained from the first $N_b + 1$ elements of the $IDFT(\tilde{\mathbf{z}} \odot \tilde{\mathbf{q}})$. Consequently, Fourier transform of the vector $\mathbf{R}_\Delta \mathbf{q}$ can be obtained by performing DFT operation on the vector $\mathbf{s}$. These two properties along with Parsevalprovide us an efficient algorithm as we will explainin the subsequent section.

3.1. Fast Algorithm

- Initialization:
  Starting from an aribitrary normalized vector $\tilde{\mathbf{b}}^0$, compute the minimum eigenvalue estimate, residual error and descent direction according to

$$\begin{aligned} \lambda^0 &= \sum_{k=0}^{N_f-1} \tilde{\mathbf{z}}[k]\, \tilde{\mathbf{v}}^0[k] \\ \tilde{\mathbf{r}}^0 &= \lambda^0 \tilde{\mathbf{b}}^0 - \tilde{\mathbf{s}}^0 \quad \tilde{\mathbf{p}}^0 = \tilde{\mathbf{r}}^0 \quad \tilde{\mathbf{v}}^0[k] = |\tilde{\mathbf{b}}^0[k]|^2 \end{aligned}$$

- Iteration:
  For $i = 0 \cdots N_b$ compute the TIR frequency response as

$$\begin{aligned} \tilde{\mathbf{b}}^{i+1} &= \tilde{\mathbf{b}}^i + \mu^i \tilde{\mathbf{p}}^i \qquad \mu^i = \frac{-B + \sqrt{B^2 - 4CD}}{2D} \\ D &= \rho_b^i \rho_c^i - \rho_a^i \rho_d^i \ , \ B = \rho_b^i - \lambda^i \rho_d^i \ , \ \ C = \rho_a^i - \lambda^i \rho_c^i \\ \rho_a{}^i &= \sum_{k=0}^{N_f-1} \tilde{\mathbf{z}}[k]\, \tilde{\mathbf{d}}_1[k] \quad , \quad \rho_b{}^i = \sum_{k=0}^{N_f-1} \tilde{\mathbf{z}}[k]\, \tilde{\mathbf{d}}_2[k] \\ \rho_c{}^i &= \sum_{k=0}^{N_f-1} \tilde{\mathbf{d}}_1[k] \quad , \quad \rho_d{}^i = \sum_{k=0}^{N_f-1} \tilde{\mathbf{d}}_2[k] \\ \tilde{\mathbf{d}}_1[k] &= (\tilde{\mathbf{p}}^i[k])^* \tilde{\mathbf{b}}^i[k] \quad , \quad \tilde{\mathbf{d}}_2[k] = (\tilde{\mathbf{p}}^{\mathbf{i}}[k])^* \tilde{\mathbf{p}}^i[k] \end{aligned}$$

Compute the minimum eigenvalue estimate, residual error, descent direction and normalized TIR vector according to

$$\begin{aligned} \lambda^{i+1} &= \frac{1}{\tau^{i+1}} \sum_{k=0}^{N_f-1} \tilde{\mathbf{z}}[k]\, \tilde{\mathbf{v}}^{\mathbf{i}}[k] \ , \quad \tilde{\mathbf{r}}^{i+1} = \frac{\lambda^{i+1} \tilde{\mathbf{b}}^{i+1} - \tilde{\mathbf{s}}^{i+1}}{\tau^{i+1}} \\ \tilde{\mathbf{p}}^{i+1} &= \tilde{\mathbf{r}}^{i+1} + \beta^i \tilde{\mathbf{p}}^i \quad , \quad \tilde{\mathbf{b}}^{i+1} = \frac{\tilde{\mathbf{b}}^{i+1}}{\tau^{i+1}} \end{aligned}$$

where

$$\beta^i = \frac{\sum_{k=0}^{N_f-1} \tilde{\mathbf{z}}[k] \ \ (\tilde{\mathbf{r}}^{i+1}[k])^* \tilde{\mathbf{p}}^i[k] + (||r^{i+1}||^2)(\rho_c{}^i + \mu^i \rho_d{}^i)}{\rho_b{}^i - \lambda^{i+1}\, \rho_d{}^i}$$

$$\begin{aligned} \tilde{\mathbf{v}}^{i+1}[k] &= \frac{\tilde{\mathbf{v}}^i[k]}{(\tau^i)^2} + (\mu^i)^2 \tilde{\mathbf{d}}_2[k] + 2\mu^i \Re(\tilde{\mathbf{d}}_1[k]) \\ \tau^{i+1} &= \sum_{k=0}^{N_f-1} (\tilde{\mathbf{v}}^{i+1}[k]) \end{aligned}$$

Upon computing the optimum setting for TIR, TEQ's coefficients are obtained from equation (2). It is also worthwhile to remark that the term $\sigma_x^2 \mathbf{H}_\Delta \mathbf{b}$ in equation (2) can be computed efficiently using equation (6). Table 1 compares the computational complexity of the proposed method against the standard matrix inversion method.

	Power iteration	Proposed algorithm
Hessian matrix Computation	$N_f^2(N_b+1)$ $+N_f(N_b+1)^2$	——
Min. eigen. Computation	$O((N_b+1)^3)$	$O(N_f log_2 N_f)$ per iteration
Sensitivity to condition No.	Highly sensitive	Robust
Other Features	Requires matrix inversion	Parallelizable, requires large N_f

Table 1: Comparison Between Proposed Method and Standard Power iteration Algorithm

4. SIMULATIONS AND PERFORMANCE EVALUATION OF THE ALGORITHM

In this section we apply the proposed algorithm for equalization of DMT in ADSL environment. A series of simulations are performed on 2 kft, 26 gauage (AWG) wire line sampled at 2.208 MHz. The power spectral density of near-end crosstalk (NEXT) noise is generated by exciting the NEXT coupling filter $|H_x(f)|^2 = k_{NEXT} f^{3/2}$ by a white Gaussian noise with power of 10mW. Unless specified, k_{NEXT} is fixed to 10^{-13}. Also there is an AWGN with power of -30dBm across the two sided spectral bandwidth. Decision delay is set to the optimum delay obtained from MMSE-UEC .Unless specified, transmit power is set such that the matched filter bound ($\mathbf{MFB} = ||h||^2 \sigma_x^2/\sigma_n^2$) of 15 dB is achieved at the receiving point. As a performance measure, we compute signal power to MSE ($SNR = \sigma_x^2/\mathbf{b}^* \mathbf{R}_\Delta \mathbf{b}$) to evaluate the performance of the methods. Throughout the simulations, performance of the proposed method is compared against MMSE-UEC method. The following points can be inferred from the plots.

- The gap between exact solution and proposed algorithm reduces as the length of the TIR filter increases (Figure 2). This is due to the fact that the Rayleigh minimization algorithm provides more exact solutions as dimension of the Hessian matrix increases.
- As long as N_f is large enough to satisfy the asymptotic equivalence of Toeplitz and circulant matrices, the proposed algorithm provides a robust solution for various values of N_f.
- Performance of the algorithm is not influenced by the spectrum of the noise. As is shown in Figure (4) Signal to MSE is constant over a large range of K_{NEXT}.
- Signal to MSE is a linear function of signal power (MFB). This is a favorable characteristic, as there would be no limitation on the dynamic range of transmit power.

5. CONCLUDING REMARKS

We have developed a novel fast algorithm as a straightforward application of Rayleigh minimization approach to solving the optimum MMSE-UEC equalization problem. Its structure was chosen to allow the use of the DFT operation which makes the algorithm highly parallelizable. The proposed method can be customized to provide a balance between performance and computational complexity. Simulation results in this paper show that the numerical

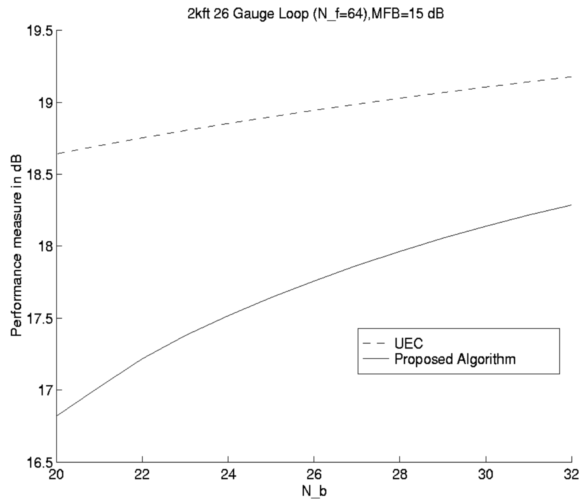

Figure 2: Performance with Different TIR Lengths

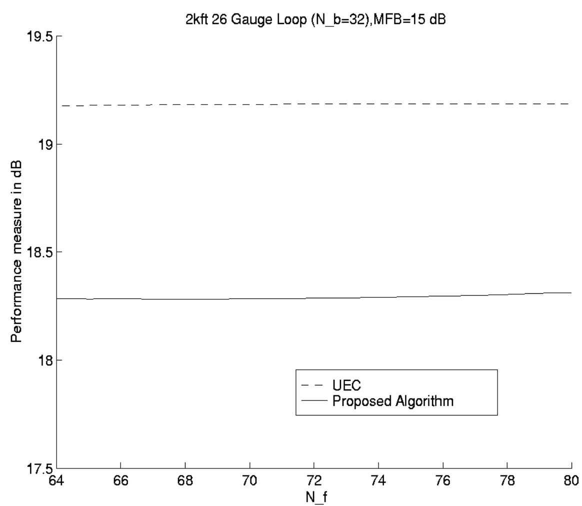

Figure 3: Performance with Different TEQ Lengths

complexity in the minimum eigenvector estimation can be reduced considerably by exploiting the proposed algorithm, without significant loss in the performance.

6. REFERENCES

[1] N. Al-Dhahir, J. Cioffi "Efficiently Computed Reduced-Parameter Input-Aided MMSE equalizers for ML Detection: A Unified Approach", *IEEE Trans. on Inform. Theory*, vol. IT-42, No. 3., pp. 903-915, May 1996.

[2] P. A. Regalia "An Unbiased Equation Error Identifier And Reduced-Order Approximations" *IEEE Trans. on Signal Processing*, vol. 42, No. 6., pp. 1397-1412, May 1994.

[3] N. Al-Dhahir, J. Cioffi "Stable Pole-Zero Modeling of Long FIR Filters With Application to MMSE-DFE", *IEEE Trans. on Communications*, vol. 45, No. 5., pp. 508-513, May 1997.

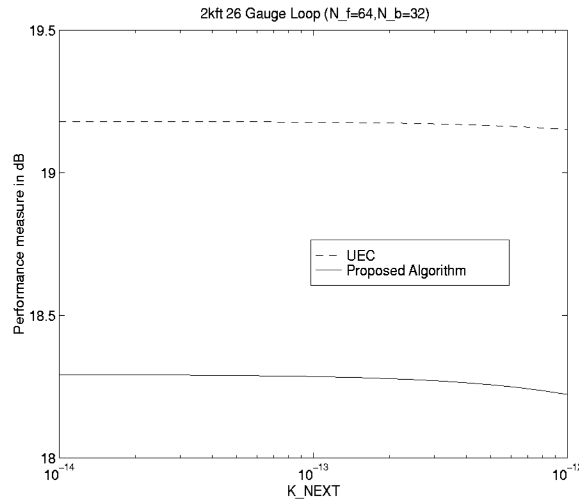

Figure 4: Performance as a Function of $K_N EXT$

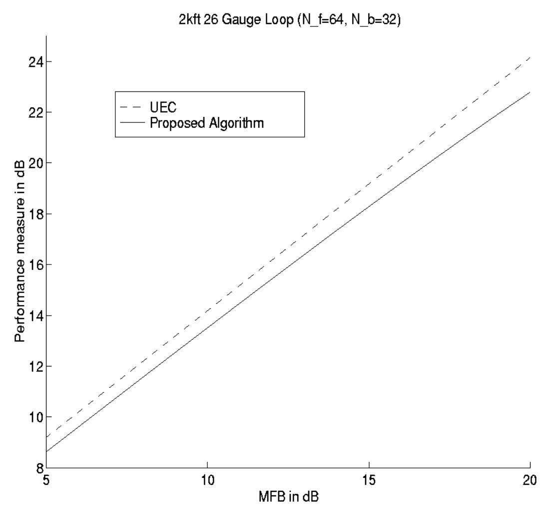

Figure 5: Performance as a Function of MFB

[4] I. Lee, J. Cioffi "A Fast Computation Algorithm for The Decision Feedback Equalizer", *IEEE Transactions on Communications*, vol. 43, No. 11., pp. 2742-2749, Nov. 1995.

[5] X. Yang, T. Sarkar, E. Arvas "A Survey of Conjugate Gradient Algorithms for Solution of Extreme Eigen-problems of a Symmetric Matrix", *IEEE Trans. on Acoustics, Speech and Signal Processing*, vol. 37, No. 10, pp. 1550-1555, Oct. 1989.

[6] T. Sarkar ,X. Yang "Application of the Conjugate Gradient and Steepest Descent for Computing the Eigenvalues of an Operator ", *Signal Processing*, vol. 17, pp. 31-38, 1989.

[7] Z. Fu, E. Dowling "Conjugate Gradient Eigenstructure Tracking for Adaptive Spectral Estimation" *IEEE trans. Signal Processing*, vol. 43, No. 3,pp. 1151-1160, MAY 1995.