Modified Leaky LMS Algorithm for Channel Estimation in DS-CDMA Systems

Sungkwon Jo†, Jihoon Choi‡, and Yong H. Lee‡

† Samsung Electronics Co. Ltd., Korea

‡ Dept. of Electrical Engineering and Computer Science, KAIST, Korea

Email: yohlee@ee.kaist.ac.kr

Abstract— **A simple adaptive least mean square (LMS) type channel estimation is developed through some modification of FIR Wiener filtering. A condition that guarantees the convergence of the algorithm and theoretical mean square error (MSE) values are derived. Computer simulation results demonstrate that the proposed algorithm can yield smaller MSE than existing techniques, and that its performance is close to that of optimal Wiener filtering.**

I. Introduction

In direct sequence code division multiple access (DS-CDMA) systems, a pilot channel is usually employed for synchronization and channel estimation. After despreading a pilot channel at a finger of a DS-CDMA RAKE receiver, the received signal $x(k)$ can be expressed as

$$x(k) = h(k) + n(k) \tag{1}$$

where $h(k)$ is a wide-sense stationary channel parameter and $n(k)$ is zero-mean additive white noise. Estimating $h(k)$ from $\{x(k)\}$ is not a trivial task, because the statistics of $h(k)$ are unknown and vary depending on the velocity of a mobile receiver. Use of an adaptive filter for this estimation is also difficult, since it is impossible to provide a training sequence consisting of channel parameters. A popular approach to the channel estimation is to use a fixed lowpass filter (LPF) whose cutoff frequency is set to the maximum Doppler frequency [1], [2]. Such an LPF is simple to implement, but tends to exhibit poor performance when its cutoff frequency differs from the actual maximum Doppler frequency. Use of optimal filters such as Wiener and Kalman filters has been proposed in [3]–[5]. These filters require channel statistics, and thus their implementation is difficult. In [6], channel parameters are estimated by applying adaptive linear prediction. The adaptive predictor can perform better than LPFs in slow fading environments, but its performance is degraded rather rapidly as the channel fade rate increases.

In this paper, we develop a new adaptive filter for channel estimation by modifying the FIR Wiener filtering formulation. The proposed adaptation scheme resembles the leaky LMS algorithm [7]. It will be shown that the performance of the adaptive estimator is close to that of the optimal Wiener estimator. The proposed technique is a useful alternative to existing channel estimates such as an LPF and an adaptive predictor.

II. Derivation of the Proposed Channel Estimation Algorithm

Suppose that an estimate of $h(k)$ in (1) is given by

$$\hat{h}(k) = \mathbf{w}^H \mathbf{x}(k) \tag{2}$$

where $\mathbf{w}$ is an N-dimensional tap weight vector and $\mathbf{x}(k) = [x(k), x(k-1), \cdots, x(k-N+1)]^T$. The optimal tap weight $\mathbf{w}_o$ minimizing $E[|h(k) - \hat{h}(k)|^2]$ satisfies the Wiener-Hopf equation,

$$\mathbf{R}\mathbf{w}_o = \mathbf{p}_{xh} \tag{3}$$

where $\mathbf{R} = E[\mathbf{x}(k)\mathbf{x}^H(k)]$ and $\mathbf{p}_{xh} = E[\mathbf{x}(k)h^*(k)]$. When $\{h(k)\}$ and $\{n(k)\}$ are uncorrelated, $\mathbf{p}_{xh}$ is rewritten as

$$\mathbf{p}_{xh} = \mathbf{p}_{xx} - [\sigma^2, 0, \cdots, 0]^T \tag{4}$$

where $\mathbf{p}_{xx} = E[\mathbf{x}(k)x^*(k)]$ and σ^2 is the variance of $n(k)$. Using (4) in (3), we get

$$(\mathbf{R} + \mathbf{A})\mathbf{w}_o = \mathbf{p}_{xx} \tag{5}$$

where $\mathbf{A}$ is an $N \times N$ matrix consisting of all zeros with the exception that (1,1)th element, say α_o, is nonzero. This parameter α_o is given by

$$\alpha_o = \frac{\sigma^2}{w_{o,0}} \tag{6}$$

where $w_{o,0}$ is the first element of the optimal tap weight vector $\mathbf{w}_o$. Throughout this paper, it is assumed that σ^2 is known[1]. The equation (5) can be viewed as the Wiener-Hopf equation associated with the correlation matrix $(\mathbf{R} + \mathbf{A})$ and the correlation vector $\mathbf{p}_{xx}$. It can be derived by minimizing $E[|x(k) - \mathbf{w}^H\mathbf{x}(k)|^2 + \mathbf{w}^H\mathbf{A}\mathbf{w}]$, and the gradient of this cost is:

$$\begin{aligned} &2\{-\mathbf{p}_{xx} + (\mathbf{R} + \mathbf{A})\mathbf{w}\} \\ &\quad = 2E[-\mathbf{x}(k)x^*(k) + (\mathbf{x}(k)\mathbf{x}^H(k) + \mathbf{A})\mathbf{w}]. \end{aligned} \tag{7}$$

Therefore, the steepest descent algorithm for obtaining $\mathbf{w}_o$ in (5) is given by

$$\begin{aligned} \mathbf{w}(k+1) &= \mathbf{w}(k) - \mu E[-\mathbf{x}(k)x^*(k) \\ &\quad + (\mathbf{x}(k)\mathbf{x}^H(k) + \mathbf{A})\mathbf{w}]. \end{aligned} \tag{8}$$

After removing the expectation $E[\cdot]$ in (8), the following least mean square (LMS) algorithm is derived:

$$\begin{aligned} \mathbf{w}(k+1) &= (\mathbf{I} - \mu\mathbf{A})\mathbf{w}(k) \\ &\quad + \mu\mathbf{x}(k)\{x(k) - \mathbf{w}^H(k)\mathbf{x}(k)\}^*. \end{aligned} \tag{9}$$

The weight $\mathbf{w}(k)$ in (9) approaches $\mathbf{w}_o$ in (5) (or (3)) as k increases. This recursion does not require any training sequence, but it needs the knowledge about α_o (or equivalently $\mathbf{A}$). In the proposed algorithm, α_o is also updated at each recursion: it is denoted by $\alpha(k)$ at the k-th iteration and updated by

$$\alpha(k) = \frac{\sigma^2}{w_0(k)} \tag{10}$$

where $w_0(k)$ is the first element of $\mathbf{w}(k)$ (compare (10) with (6)). The proposed channel estimation algorithm is summarized as follows:

1. Channel estimate:

$$\hat{h}(k) = \mathbf{w}^H(k)\mathbf{x}(k) \tag{11}$$

2. Estimation error:

$$e(k) = x(k) - \hat{h}(k) \tag{12}$$

3. Tap-weight adaptation:

$$\mathbf{w}(k+1) = (\mathbf{I} - \mu\hat{\mathbf{A}}(k))\mathbf{w}(k) + \mu\mathbf{x}(k)e^*(k) \tag{13}$$

where $\hat{\mathbf{A}}(k)$ is an $N \times N$ matrix consisting of all zeros with the exception that the (1,1)th element is $\alpha(k)$ in (10). The proposed algorithm is referred to as the *modified leaky* LMS algorithm, due to its resemblance to the leaky LMS algorithm, and the parameter α_o is called the optimal *leakage* coefficient.

III. Performance Analysis

In this section, the stability condition and the MSE associated with the recursive algorithm in (13) are derived. Then the results are confirmed through computer simulation.

A. Theoretical Derivation

From (10), the recursive equation in (13) is rewritten as

$$\mathbf{w}(k+1) = \mathbf{w}(k) + \mu\mathbf{x}(k)\{x(k) - \mathbf{w}^H(k)\mathbf{x}(k)\}^* - \mu\mathbf{c} \tag{14}$$

where $\mathbf{c} = [\sigma^2, 0, \cdots, 0]^T$. Subtracting the optimal tap weight $\mathbf{w}_o$ from both sides of (14), we get

$$\begin{aligned} \boldsymbol{\epsilon}(k+1) &= (\mathbf{I} - \mu\mathbf{x}(k)\mathbf{x}^H(k))\boldsymbol{\epsilon}(k) \\ &\quad + \mu\{\mathbf{x}(k)(x(k) - \mathbf{w}_o^H(k)\mathbf{x}(k))^* - \mathbf{c}\} \\ &= (\mathbf{I} - \mu\mathbf{x}(k)\mathbf{x}^H(k))\boldsymbol{\epsilon}(k) \\ &\quad + \mu\{\mathbf{x}(k)(e_o(k) + n(k))^* - \mathbf{c}\} \end{aligned} \tag{15}$$

[1] σ^2 can be easily estimated by measuring the power of the signal which is despread with a code that is orthogonal to all codes which are being used [8].

where $\boldsymbol{\epsilon}(k) = \mathbf{w}(k) - \mathbf{w}_o$ and $e_o(k) = h(k) - \mathbf{w}_o^H(k)\mathbf{x}(k)$. Assuming that $\boldsymbol{\epsilon}(k)$ and $\mathbf{x}(k)$ are uncorrelated and that $E[\boldsymbol{\epsilon}(k)] = 0$, the correlation matrix for $\boldsymbol{\epsilon}(k)$, $\mathbf{K}(k) = E[\boldsymbol{\epsilon}(k)\boldsymbol{\epsilon}^H(k)]$, can be computed as follows.

$$\begin{aligned}\mathbf{K}(k+1) &= (\mathbf{I} - \mu\mathbf{R})\mathbf{K}(k)(\mathbf{I} - \mu\mathbf{R}) \\ &\quad + \mu^2\{(J_{min} + \sigma^2)\mathbf{R} + \mathbf{c}\mathbf{c}^H\} \qquad (16)\end{aligned}$$

where $J_{min} = E[|e_o(k)|^2]$, denotes the MSE of the Wiener filter. The estimation error can be expressed as

$$\begin{aligned}e(k) &= h(k) - \mathbf{w}^H\mathbf{x}(k) \\ &= h(k) - \mathbf{w}_o^H\mathbf{x}(k) - \boldsymbol{\epsilon}^H(k)\mathbf{x}(k) \\ &= e_o(k) - \boldsymbol{\epsilon}^H(k)\mathbf{x}(k). \qquad (17)\end{aligned}$$

Under the assumption that $e_o(k)$ and $\boldsymbol{\epsilon}^H(k)\mathbf{x}(k)$ are independent, the mean squared error (MSE) at iteration k, say $J(k)$, can be represented as

$$\begin{aligned}J(k) &= E[|e(k)|^2] \\ &= J_{min} + E[\Delta\mathbf{w}^H(k)\mathbf{x}(k)\mathbf{x}^H(k)\Delta\mathbf{w}(k)] \\ &= J_{min} + tr[\mathbf{R}\mathbf{K}(k)] \qquad (18)\end{aligned}$$

where $tr[\cdot]$ means a trace operation. Following the approach in [7, p.397], it can be shown that $J(k)$ converges to a constant if and only if

$$0 < \mu < \frac{2}{\lambda_{max}} \qquad (19)$$

where λ_{max} is the maximum eigenvalue of $\mathbf{R}$. When (19) is satisfied, the MSE is given by

$$J(k) = J_{min} + \boldsymbol{\lambda}^T\{\mathbf{B}^k(\mathbf{x}(0) - \mathbf{x}(\infty)) + \mathbf{x}(\infty)\} \qquad (20)$$

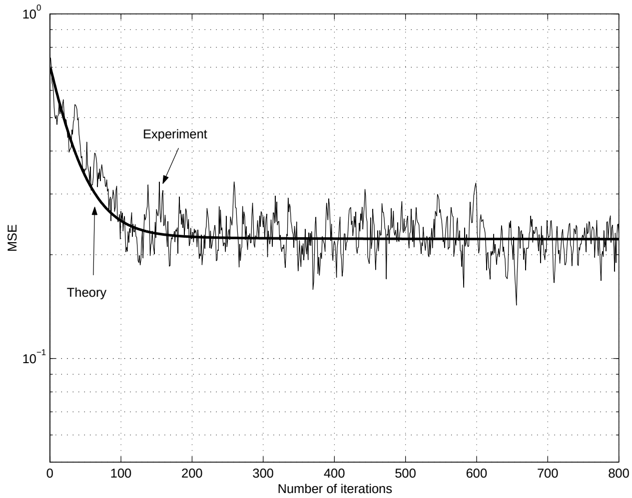
where the various terms are defined as follows:

- $\boldsymbol{\lambda}$ is an $N \times 1$ vector whose elements are the eigenvalues of $\mathbf{R}$.
- $\mathbf{B}$ is an $N \times N$ matrix with elements $b_{ij} = \begin{cases} (1-\mu\lambda_i)^2, & i = j \\ \mu^2\lambda_i\lambda_j, & i \neq j \end{cases}$.
- $\mathbf{x}(0) = diag\{\mathbf{Q}^H(\mathbf{w}(0) - \mathbf{w}_o)(\mathbf{w}(0) - \mathbf{w}_o)^H\mathbf{Q}\}$ when $\mathbf{Q}$ is the unitary matrix consisting of the eigenvectors for $\mathbf{R}$.
- $\mathbf{x}(\infty) = \mu^2(\mathbf{I} - \mathbf{B})^{-1}\{(J_{min} + \sigma^2)\boldsymbol{\lambda} + [\sigma^4, 0, \cdots, 0]^T\}$.

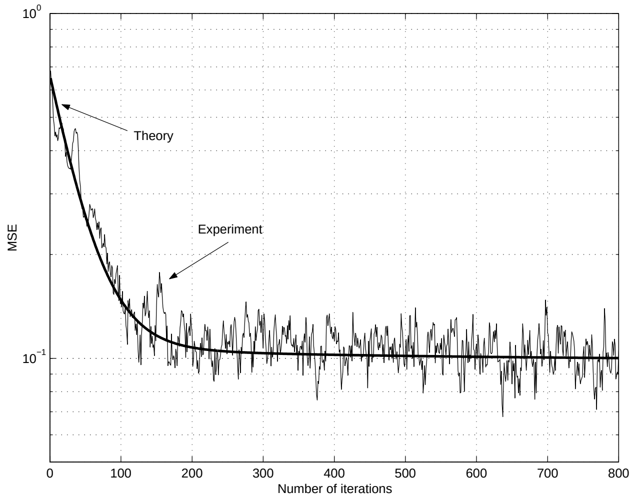
The steady-state MSE is written as

$$\begin{aligned}J(\infty) &= J_{min} + \boldsymbol{\lambda}^T\mu^2(\mathbf{I} - \mathbf{B})^{-1} \\ &\quad \cdot\{(J_{min} + \sigma^2)\boldsymbol{\lambda} + [\sigma^4, 0, \cdots, 0]^T\}. \qquad (21)\end{aligned}$$

B. Comparison of Experimental Results with Theory



(a)



(b)

Fig. 1. Comparison of experimental results with theory for the proposed method when $f_dT = 0.03$, $N = 5$, and $\mu = 0.0025$. (a) SNR= 0 dB (b) SNR= 5 dB.

To confirm the theoretical result in (20), the modified leaky LMS algorithm was applied to channel estimation and the empirical MSEs were obtained. The signal $x(k)$ in (1) was generated under the following assumptions: $h(k)$ is a Rayleigh fading gain with $f_dT = 0.03$ and $n(k)$ is a complex additive white Gaussian noise (AWGN), where f_dT is the normalized maximum Doppler frequency[2]. The parameters of the modified leaky LMS algorithm were as follows: the number of taps $N = 5$; the initial tap weight vector $\mathbf{w}(0) = [1/N, 0, \cdots, 0]^T$; and the step-size $\mu = 0.0025$. Fig. 1 shows the learning curves when SNR= 0 dB and SNR= 5 dB. The experimental curves were obtained by averaging the squared error over 100 independent trials. The results demonstrate a good agreement between theory and experiment.

IV. Application to Channel Estimation

The signal $x(k)$ in (1) was generated as in Section III.B, with the exception that various f_dT values between 0 and 0.05 were considered. The parameters of the proposed algorithm remained the same. For comparison, two LPFs which were designed for $f_dT = 0.03$ and 0.05, the adaptive linear predictor, and the optimal Wiener filter were considered as well as the proposed method. The parameters of these filters were the same as the corresponding parameters of the modified leaky LMS algorithm. The LPFs were 5-tap FIR filters which were designed using the Parks-McClellan algorithm. The Wiener filter was designed for each f_dT, and thus it gave a minimum bound. The empirical MSEs of the proposed and the adaptive linear predictor were obtained by averaging the steady state squared error values.

Fig. 2 compares the empirical MSE values. It was seen that the performance of the proposed algorithm was close to that of the Wiener filter. The proposed method always outperformed the adaptive linear predictor, and it almost always performed better than the LPF: in Fig. 2(a), the LPF designed for $f_dT = 0.03$ exhibited slightly smaller MSEs than the proposed in the vicinity of $f_dT = 0.03$; but for the other f_dT values, the latter yielded smaller MSEs than the former. The adaptive linear predictor worked better than the LPFs for slow fading channels, but its MSE rapidly increased as f_dT increased. As expected, the LPF performed well

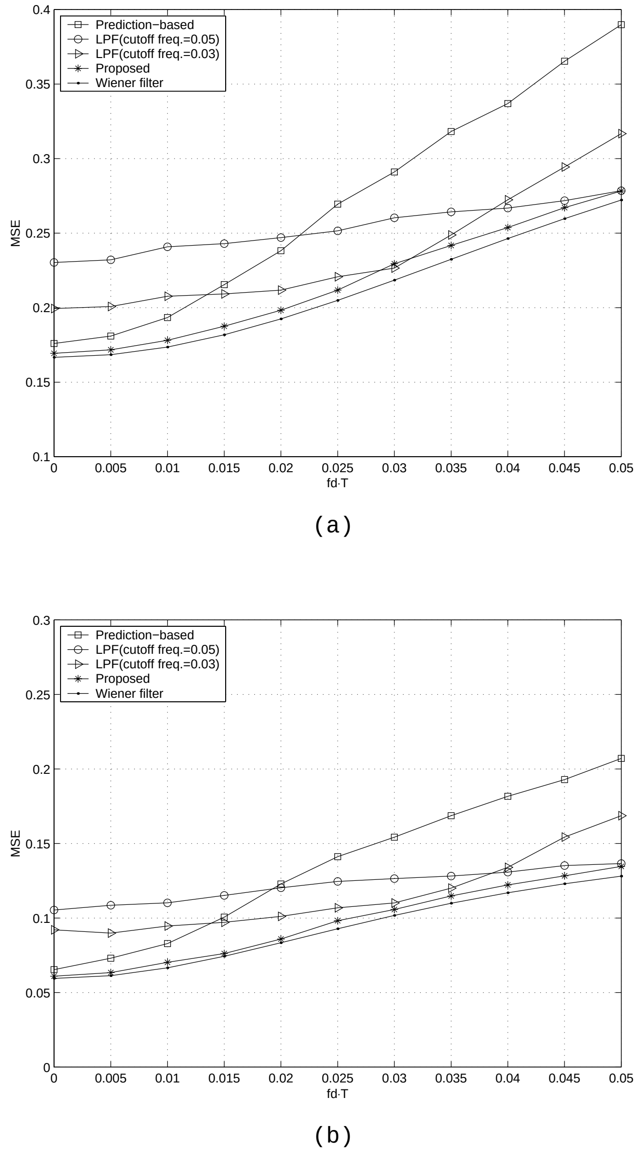


Fig. 2. MSE comparison when $N = 5$ and $\mu = 0.0025$. (a) SNR= 0 dB (b) SNR= 5 dB.

[2]In the case of W-CDMA, f_dT is 0.037 when the chip rate is 3.84 Mcps, the symbol rate is 15.0 ksps, the carrier frequency is 2.0 GHz, and the maximum mobile velocity is 300 km/h.

only in the neighborhood of the $f_d T$ value for which it was designed. The modified leaky LMS algorithm should be a useful alternative to the LPF and adaptive linear predictor.

References

[1] M. L. Moher and J. H. Lodge, "TCMP-A modulation and coding strategy for Rician fading channels," *IEEE J. Select. Areas Commun.*, vol. SAC–7, pp. 1347–1355, Dec. 1989.

[2] S. Sampei and T. Sunaga, "Rayleigh fading compensation method for 16QAM in digital land mobile radio channels," *Proc. VTC'89*, pp. 640–646, May 1989, San Francisco, CA, Dec. 1989.

[3] J. K. Cavers, "An analysis of pilot symbol assisted modulation for Rayleigh fading channels," *IEEE Trans. Veh. Technol.*, vol. VT–40, pp. 686–693, Nov. 1991.

[4] A. N. D'Andrea and A. Diglio and U. Mengali, "Symbol-aided channel estimation with nonselective Rayleigh fading channels," *IEEE Trans. Veh. Technol.*, vol. VT–44, pp. 41–49, Feb. 1995.

[5] A. Aghamohammadi and H. Meyr and G. Ascheid, "Adaptive synchronization and channel parameter estimation using an extended Kalman filter," *IEEE Trans. Commun.*, vol. COM–37, pp. 1212–1218, Nov. 1989.

[6] Y. Liu and S. D. Blostein, "Identification of frequency nonselective fading channels using decision feedback and adaptive linear prediction," *IEEE Trans. Commun.*, vol. COM-43, pp. 1484–1492, Feb./Mar./Apr. 1995.

[7] S. Haykin, *Adaptive Filter Theory*, Prentice-Hall, 1996, 3rd ed.

[8] S. Ulukus and R. D. Yates, "A blind adaptive decorrelating detector for CDMA systems," *IEEE J. Select. Areas Commun.*, vol. 16, pp. 1530–1541, Oct. 1998.