2001年度 第14回 信號處理合同學術大會 論文集 第14卷 1號

A New Convergence Behavior of the Least Mean *K*-power Adaptive Algorithm

Kang Seung Lee

Dept. of Computer Eng., Dong-Eui University
E-mail : kslee@dongeui.ac.kr

Abstract

In this paper we study a new convergence behavior of the least mean fourth (LMF) algorithm where the error raised to the power of four is minimized for a multiple sinusoidal input and Gaussian measurement noise. Here we newly obtain the convergence equation for the sum of the mean of the squared weight errors, which indicates that the transient behavior can differ depending on the relative sizes of the Gaussian noise and the convergence constant. It should be noted that no similar results can be expected from the previous analysis by Walach and Widrow[1].

I. Introduction

The Least Mean Square(LMS) adaptive algorithm have been successfully utilized for a variety of applications including system identification[2,3,4], noise cancellation[5,6], echo cancellation[7,8], channel equalization[9] during the last two decades. Meanwhile, the adaptive filtering algorithms that are based on high order error power conditions have been proposed and their performances have been investigated[1,10,11,12,13]. Despite the potential advantages, these algorithms are less popular than the conventional LMS algorithm in practice. This seems partly because the analysis of the high order error based algorithms is much more difficult, and thus not much still has been known about the algorithms.

The least mean fourth (LMF) adaptive algorithm[1] in which the error raised to the power of four is minimized. Here, one has to consider the possibility of the convergence to local minimum. However, the mean of the error to the power of four is a convex function of the weight vector and therefore can not have local minima. Indeed the Hessian matrix of the error to the four power function can be shown to be positive define or positive semidefinite[14].

Walach and Widrow studied the convergence of the least mean fourth (LMF) adaptive algorithm[1]. However, in their convergence study of the mean squared weight errors, the statistical moments of the weight errors with the orders greater than two were neglected and the transient behavior was not analyzed. In this paper, we present a new result on the convergence of the least mean fourth algorithm under the system identification model with the multiple sinusoidal input and Gaussian measurement noise.

Following the introduction, we give a brief description of the underlying system model in Section II. The results of the convergence analysis and the simulation are presented in Sections III and IV, respectively. Finally we make a conclusion in Section

II. System Model

We consider an adaptive noise cancellation problem for the multiple sinusoidal input and Gaussian measurement noise. In that case, both the unknown system and corresponding adaptive filter can be described by the multiple in-phase(I) and quadrature(Q) weights as shown in Figure 1[2,5].

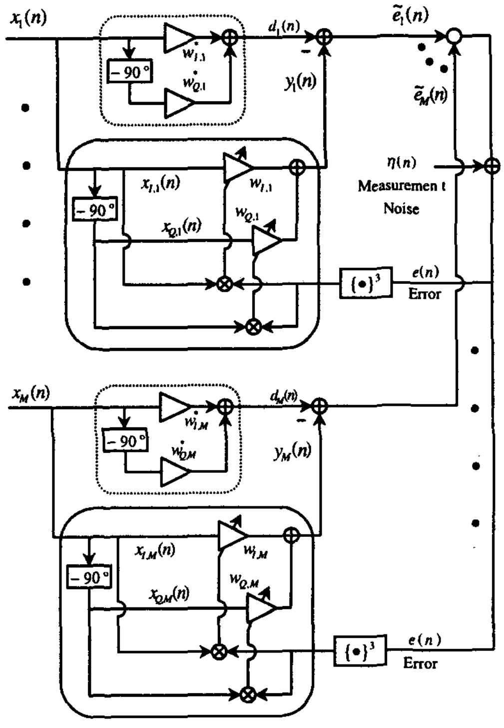

Figure 1. Adaptive digital filter for a multiple sinusoidal input under study.

For the m-th sinusoidal noise, the adaptive canceller structure also becomes to have two weights $w_{I,m}(n)$ and $w_{Q,m}(n)$, with I and Q inputs, $x_{I,m}(n)$ and $x_{Q,m}(n)$, respectively. Thus the output of the m-th controller, $y(n)$ is expressed as

$$y(n) = \sum_{m=1}^{M} \left\{ w_{I,m}(n) x_{I,m}(n) + w_{Q,m}(n) x_{Q,m}(n) \right\} \tag{1}$$

where

$$x_{I,m}(n) = A_m \cos(\omega_m n + \phi_m) \triangleq A_m \cos \Psi_m(n),$$

$$x_{Q,m}(n) = A_m \sin(\omega_m n + \phi_m) \triangleq A_m \sin \Psi_m(n),$$

m : branch index = 1, 2, 3, ..., M,

$x_{Q,m}(n) = A_m \sin(\omega_m n + \phi_m) \triangleq A_m \sin\Psi_m(n)$,

m : branch index = 1, 2, 3, ..., M,

n : discrete time index,

A : amplitude,

ω: normalized frequency,

: random phase.

Also, referring to the notation in Figure. 1, the error signal $e(n)$ is represented by

$$\begin{aligned} e(n) &= \sum_{m=1}^{M}\{d_m(n) - y_m(n)\} + \eta(n) \\ &= -\sum_{m=1}^{M} A_m[\{w_{I,m}(n) - w_{I,m}^*\}\cos\Psi_m(n) + \{w_{Q,m}(n) - w_{Q,m}^*\}\sin\Psi_m(n)] \\ &\quad + \eta(n) \end{aligned} \tag{2}$$

where $\eta(n)$ is zero-mean measurement noise.

It can be shown from (1) and (2) that minimizing the fourth power error and using a gradient-descent method[2] yields a pair of the LMF weight update equations for each m as

$$w_{I,m}(n+1) = w_{I,m}(n) + 2\mu_m e^3(n) x_{I,m}(n),$$

and $$w_{Q,m}(n+1) = w_{Q,m}(n) + 2\mu_m e^3(n) x_{Q,m}(n) \tag{3}$$

where μ_m is a convergence constant.

In the following, we analyze the convergence behavior of the summed variance of weight errors of the LMF algorithm using a new analysis method.

III. Convergence Analysis

To see how the adaptive algorithm derived in (3) converges, we first investigate the convergence of the expected values of the adaptive weights. To simplify the convergence equation, we may introduce two weight errors as

$$v_{I,m}(n) \triangleq w_{I,m}(n) - w_{I,m}^*$$

and $$v_{Q,m}(n) \triangleq w_{Q,m}(n) - w_{Q,m}^* \tag{4}$$

Inserting (4) into (3), we have

$$v_{I,m}(n+1) = v_{I,m}(n) + 2\mu_m e^3(n) x_{I,m}(n),$$

and $$v_{Q,m}(n+1) = v_{Q,m}(n) + 2\mu_m e^3(n) x_{Q,m}(n). \tag{5}$$

Next we investigate the convergence of the mean- square error(MSE), $E\left[e^2(n)\right]$ Using (2) and (4), we can express the MSE as

$$\begin{aligned} E\left[e^2(n)\right] &= \sum_{m=1}^{M} e_m^2(n) + \sigma_\eta^2 \\ &= \frac{1}{2}\sum_{m=1}^{M} A_m^2 \xi_m(n) + \sigma_\eta^2 \end{aligned} \tag{6}$$

where

$$\xi_m(n) \triangleq E\left[v_{I,m}^2(n)\right] + E\left[v_{Q,m}^2(n)\right],$$

$$\sigma_\eta^n \triangleq E\left[\eta^2(n)\right]$$

From (6), we find that studying the convergence of MSE is directly related to studying the sum of $\xi_m(n)$.

Inserting (1) and (2) into (5), and assuming that input signal $x_m(n)$ measurement noise $\eta(n)$, and weight errors $v_{I,m}(n)$, $v_{Q,m}(n)$ are independent of each other, we take the statistical average of both sides to obtain two equations for $E\left[v_I^2(n+1)\right]$, $E\left[v_Q^2(n+1)\right]$. Since there two equations are symmetrical, we add them and assume that $E\left[v_{I,m}^2(n+1)\right] \cong E\left[v_{Q,m}^2(n+1)\right]$. Thus, eliminating the subscripts I and Q to simplify the second moment equation of weight error and rearranging the terms yields .

$$\begin{aligned} E[v_m^2(n+1)] &= \frac{5}{4}\mu_m^2 A_m^8\{E[v_m^6(n)] + 3E[v_m^2(n)]E[v_m^4(n)]\} \\ &\quad - \frac{3}{2}\mu_m A_m^4\{E[v_m^4(n)] + (E[v_m^2(n)])^2\} \\ &\quad + \frac{45}{2}\mu_m^2 A_m^6 E[\eta^2(n)]\{E[v_m^4(n)] + (E[v_m^2(n)])^2\} \\ &\quad + \{1 - 6\mu_m A_m^2 E[\eta^2(n)] + 30\mu_m^2 A_m^4 E[\eta^4(n)]\}E[v_m^2(n)] \\ &\quad + 2\mu_m^2 A_m^2 E[\eta^6(n)]. \end{aligned} \tag{7}$$

Assuming that $\eta(n)$ is a Gaussian with zero average and $w_{I,m}(n)$, $w_{Q,m}(n)$ are Gaussian variables, $v_m(n)$ is also a Gaussian variable. Thus, (7) can be simplified by expressing $E\left[v_m^{2K}(n)\right]$ as $E\left[v_m^2(n)\right]$. Although $\eta(n)$, $E\left[v_m(n)\right]$ decreases very rapidly, it is no zero from the beginning. Thus, a Gaussian random variable $\Delta w_m(n)$ with zero average, and its variance are adapted as follows:

$$\Delta w_m(n) \triangleq v_m(n) - V_m(n),$$

and $$E\left[v_m^2(n)\right] = V_m^2(n) + \rho_m^2(n) \tag{8}$$

where $V_m(n) \triangleq E\left[v_m(n)\right]$

$\rho_m^2(n) \triangleq E\left[\Delta^2 w_m(n)\right]$

From (8), we find that during the transient state, i.e. from beginning to the moment just before the steady state, $\rho_m^2(n)$ is much smaller than $V_m^2(n)$ and $E\left[v_m(n)\right]$ and be regarded as $V_m^2(n)$. On the other hand, $\rho_m^2(n)$ becomes dominant over $V_m^2(n)$ in the steady state and $E\left[v_m(n)\right]$ can be regarded as $\rho_m^2(n)$.

Now, we apply (8) to (7) and use the relationship between $E\left[v_m^{2K}(n)\right]$ and $E\left[v_m^2(n)\right]$ of the Gaussian random variable[15] to arrive at the following equation.

$$
\begin{aligned}
&V_m^2(n+1)+\rho_m^2(n+1)\\
&=5\mu_m^2A_m^8\{V_m^6(n)+9\rho_m^2(n)V_m^4(n)+18\rho_m^4(n)V_m^2(n)+6\rho_m^6(n)\}\\
&\quad-(3\mu_mA_m^4-45\mu_m^2A_m^6\sigma_\eta^2)\{V_m^4(n)+4\rho_m^2(n)V_m^2(n)+2\rho_m^4(n)\\
&\quad+(1-6\mu_mA_m^2\sigma_\eta^2+\mu_m^2A_m^4\sigma_\eta^4)\{V_m^2(n)+\rho_m^2(n)\}\\
&\quad+30\mu_m^2A_m^2\sigma_\eta^6 \qquad (9)
\end{aligned}
$$

The convergence equation (9) may be examined for two different cases, First, $\rho_m^{2K}(n)$ and the last term of (9) can be removed for the transient state.

It is noted from the right-hand side of (10) that in extreme cases, only one of the two terms $V_m^6(n)$ or $V_m^2(n)$ for which those two terms are the same and is given by

$$V_{m,th}^2=\sqrt{\frac{1-6\mu_mA_m^2\sigma_\eta^2+90\mu_m^2A_m^4\sigma_\eta^4}{5\mu_m^2A_m^8}} \qquad (10)$$

In (9), the first term $V_m^6(n)$ acts as the dominant term when $V_m^2(n)$ is greater than $V_{m,th}^2$, If $V_m^6(n)$ $V_m^2(n)$ is smaller than $V_{m,th}^2$ then the last term becomes dominant. Figure 2 is given to illustrate in terms of the convergence constant μ_m and the variance of measurement noise σ_η^2, which of the two terms, the first term $V_m^6(n)$ and the last term $V_m^2(n)$, is dominant when $V_{m,th}^2$=0.8. Point (a) is a region in which the term $V_m^6(n)$ dominates over the other and point (b) is when $V_m^2(n)$ term is the dominant one. Therefore, the transient convergence equation (9) can be written as ;

$$
V_m^2(n+1)\cong\begin{cases}5\mu_m^2A_m^8V_m^6(n) & ,V_m^2(n)\rangle\rangle V_{m,th}^2 \quad (11a)\\ (1-6\mu_mA_m^2\sigma_\eta^2+90\mu_m^2A_m^4\sigma_\eta^4)V_m^2(n) & ,V_m^2(n)\langle\langle V_{m,th}^2 \quad (11b)\end{cases}
$$

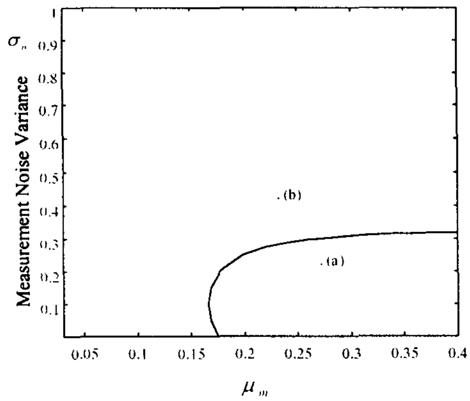

Figure 2. Dominant term decision diagram for the LMF algorithm of the summed variance of weight errors at the transient-state.
[point (a) : μ_m = 0.3 and σ_η^2 = 0.1.
point (b) : μ_m = 0.2 and σ_η^2 = 0.5.]

In the steady state, $V_m^2(n)$ becomes sufficiently small and the terms that include $\rho_m^4(n)$ and $\rho_m^6(n)$ can be ignored in the convergence equation (9). The equation is then simplified as

$$
\begin{aligned}
\rho_m^2(n+1)\cong&(1-6\mu_mA_m^2\sigma_\eta^2+90\mu_m^2A_m^4\sigma_\eta^4)\rho_m^2(n)\\
&+30\mu_m^2A_m^2\sigma_\eta^6 \qquad (12)
\end{aligned}
$$

Therefore, from (11) and (12), we can get convergence condition, time constant, an steady stare value.

IV. Computer simulations

In this section, we present the results obtained from computer simulation along with the theoretical analysis of LMF algorithm in the previous section.

case 1. The convergence property of LMF algorithm.

case 2. The performance comparison of LMF and LMS. .

We set the frequencies of the first and second sinusoidal signal at 120Hz and 240Hz, respectively, and selected 2KHz for sampling frequency. The input signal $x(n)$ and desired signal $d(n)$ are given by

$$
\begin{aligned}
x(n)&=\sum_{m=1}^{2}A_m\cos(\omega_mn+\phi_m)\\
&=\sqrt{2}\left\{\cos(\frac{240\pi n}{2000}+\phi_1)+\cos(\frac{480\pi n}{2000}+\phi_2)\right\},\\
d(n)&=\sum_{m=1}^{2}\{w_{I,m}^*x_{I,m}+w_{Q,m}^*x_{Q,m}\}\\
&=0.6x_{I,1}(n)-0.1x_{Q,1}(n)+0.3x_{I,2}(n)-0.3x_{Q,2}(n) \qquad (13)
\end{aligned}
$$

The simulation was carried out by setting 0.001 and 1 as the variances of measurement noise σ_η^2. And the initial value of weights is zero. The simulation results were obtained by ensemble averaging 1000 independent runs.

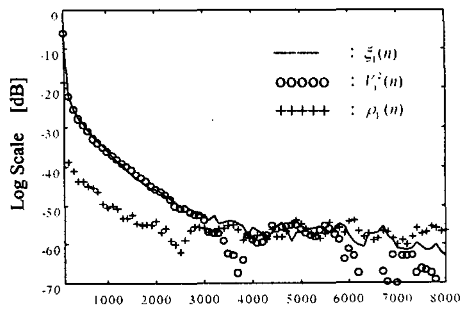

Figure 3. Learning curves for the LMF algorithm of the summed variance of weight errors when the convergence behaviors are divided between $V^2(n)$ and $\rho^2(n)$.

Figure 3 showed the summed variance convergence curves of weight error for the LMF algorithm that resulted from the simulation when $\mu_{1(LMF)} = 0.2$ $\sigma_\eta^2 = 0.001$. We see that $V^2(n)$ is the dominant term during the transient state whereas $\rho^2(n)$ becomes dominant during the steady state.

We have compared the convergence behavior of LMF algorithm and that of algorithm LMS through simulation.

In Figure 4, the convergence behavior curves of summed variance of weight error obtained from simulation are compared with each other. It has been newly found that for some region of μ and σ_η^2 , resulting in sufficiently small $V_m^2(n)$ values compared to unity, the initial convert- gence of the LMF algorithm is much faster than the conventional LMS algorithm. Later on, the LMF convergence looks similar to the LMS case. This fact has not been reported yet mainly because the higher order moments have not been included in the previous analysis of the LMF transient behavior[1]. On the other hand, when $V_{m,th}^2$ is large, the LMF algorithm converges geometrically at a rate a bit slower than the LMS case.

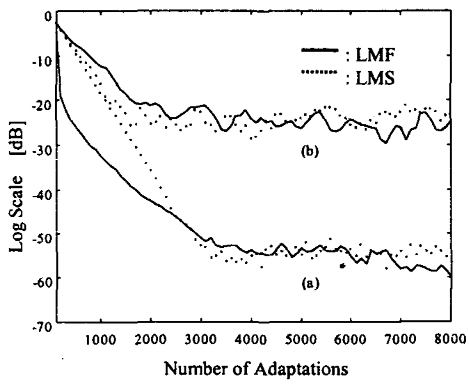

Figure 4. Comparison of the LMF and LMS algorithm learning curves of the summed variance of weight errors;

(a) $\mu_{(LMS)}$ =0.002, $\mu_{(LMF)} = 0.2$, σ_η^2 =0.001 and $V_{th}^2 = 0.558$,

(b) $\mu_{(LMS)} = 0.002$, $\mu_{(LMF)} = 0.0002$, $\sigma_\eta^2 = 1$ and $V_{th}^2 = 558$.

V. Conclusions

We present a new result on the convergence of the least mean fourth(LMF) algorithm under the system identification model with the multiple sinusoidal input and Gaussian measurement noise. The analytical result on the mean square convergence shows that depending on the power of Gaussian noise and the size of convergence factor. Accordingly, the transient behavior can be characterized by one of the two cases: (1) initially, the LMF algorithm converges much faster than the LMS, but soon after that, it converges almost linearly on logarithmic scale like the LMS algorithm; (2) the LMF algorithm converges linearly and at a slower rate than the LMS. To sum up, different convergence behavior was observed depending on the variance of Gaussian measurement noise and the magnitude of convergence constant.

References

1. E. Walach and B. Widrow, "The Least Mean Fourth (LMF) Adaptive Algorithm and Its Family," *IEEE Trans. on Information Theory*, Vol. 30, No. 2, pp. 275-283, March 1984.
2. B. Widrow and S. D. Stearns, *Adaptive Signal Processing : Prentice-Hall*, 1985.
3. C. P. Kwong, "Dual Sign Algorithm for Adaptive Filtering," *IEEE Trans. on Communications*, Vol. 34, No. 12, pp. 1272-1275, Dec. 1986.
4. S. Dasgupta and C. R. Jhonson, "Some Comments on the Behavior of Sign-sign Adaptive Identifiers," *System and Letters*, Vol. 7, pp. 75-82, April 1986.
5. B. Widrow, J. R. Glover, J. M. McCool et al., "Adaptive Noise Cancelling Principles and Applications," *Proc. IEEE*, Vol. 63, pp. 1692-1716, Dec. 1975.
6. W. A. Harrison et al., "A New Application of Adaptive Noise Cancellation," *IEEE Trans. on Acoustics, Speech, and Signal Processing*, Vol. 34, No. 1, pp. 21-27, 1986.
7. D. D. Falconer, "Adaptive Reference Echo Cancellation," *IEEE Trans. on Communications*, Vol. 30, No. 9, pp. 2083-2094, Sept. 1982.
8. A. Kanemasa and K. Niwa, "An Adaptive-step Sign Algorithm for Fast Convergence of a Data Echo Canceller," *IEEE Trans. on Communications*, Vol. 35, No. 10, pp. 1102-1108, October 1987.
9. P. F. Adam, "Adaptive Filtering in Communications", Chap. 8, *Adaptive Filters*, edited by C. F. N. Cowan and P. M. Grant, Prentice Hall, 1985.
10. S. Pei and C. Tseng, "Adaptive IIR Notch Filter Based on Least Mean p-Power Error Criterion," *IEEE Trans. On Circuits and Systems*, Vol. II-40, No. 8, pp. 525-529, aug. 1993.
11. J. Schroder, Rao Yarlagadda, and J. hershey, "Lp Normed Minimization with Applications to Linear Predictive Modeling for Sinusoidal Frequency Estimation," *Signal Processing*, Vol. 24, pp. 193-216, Aug. 1991
12. K. S. Lee and D. H. Youn, " The Filtered-x Least Mean Fourth algorithm for Active Noise Control and Its Convergence Analysis," *Jour. of the Acoustical Society of Korea*, Vol. 15, No. 3E, pp. 66-73, Sept. 1996.
13. K. S. Lee, "Performance Analysis of Adaptive Algorithms for Active Noise Control," Ph. D. Thesis, *Yonsei University*, Seoul, Aug. 1995.
14. A. Gersho, "Some Aspects of Linear Estimation with Non-Mean-Square Error Criteria," *Proc. Asilomar Ckts. and System Conf.*, 1969.
15. J. S. Bendat, *Nonlinear System Analysis and Identification from Random Data* : Jhon Wiley & Sons, 1990.