BOUNDING THE PERFORMANCE OF THE LMS ESTIMATOR FOR CASES WHERE PERFORMANCE EXCEEDS THAT OF THE FINITE WIENER FILTER

Kevin J. Quirk, James R. Zeidler, Laurence B. Milstein

Department of Electrical and Computer Engineering
University of California at San Diego, La Jolla, CA 92093 USA

ABSTRACT

The least-mean-square (LMS) estimator is a nonlinear estimator with information dependencies spanning the entire set of data fed into it. The traditional analysis techniques which are used to model this estimator obscure this, restricting the estimator to the finite set of data sufficient to span the length of its filter. The finite Wiener filter is thus often considered a bound on the performance of the LMS estimator. Several papers have reported the performance of the LMS filter exceeding that of the finite Wiener filter. In this paper, we will demonstrate a bound on the LMS estimator, which does not exclude the contributions from data outside its filter length, and which demonstrates the ability of the LMS estimator to outperform the finite Wiener filter in certain cases.

1. INTRODUCTION

The least-mean-square (LMS) adaptive filter was first introduced by Widrow. Since then, it has found widespread use in many applications [6], due in part to the simplicity of its implementation; it requires only a finite impulse response (FIR) filter and a first-order weight update equation. This simplicity, however, belies what is actually a complex nonlinear estimator. A direct analysis of this estimator's performance does not appear to be feasible; therefore, attention has focused on restricting the statistics of the input processes to simplify this analysis. The "independence assumptions" approach, outlined in Section 3 of this paper, is the most common such method. Through invocation of this strict set of assumptions, the LMS estimator is modeled as the combination of a finite Wiener filter along with some 'misadjustment noise' representing the difference between the converged LMS weights and the finite Wiener filter. This approach has been shown to be an effective way of analyzing the LMS estimator in many situations, even when the assumptions are not strictly met [7]. Based on the widespread success of this model, the finite Wiener filter is often assumed to bound the performance of LMS estimators. Recently, there have been several papers reporting cases where the performance of the LMS filter surpasses that of the finite Wiener filter [4] [5].The work in this paper continues the investigation of this phenomenon by providing a bound on the performance of the LMS estimator. In Section 4 of this paper, we will bound the performance of the LMS estimator by that of the optimal estimator for a class of signals without using the "independence assumptions." We will then, in Section 5, compare the performance of the LMS estimator with that of the optimal estimator in cases where the LMS estimator outperforms the finite Wiener filter.

This work was partially supported by the NSF Industry/University Cooperative Research Center on Integrated Circuits and Systems at UCSD, and by the Focused Research Initiative on Wireless Multimedia Networks under Grant DAA HO4-95-1-0248.

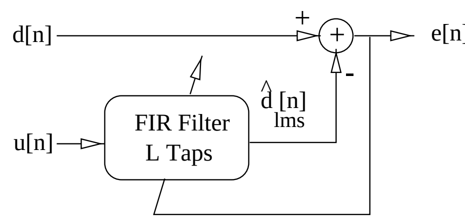

Figure 1: LMS Estimator

2. BACKGROUND

It is necessary to estimate the current value of a desired signal, $d[n]$. To form an estimate, $\hat{d}[n]$, we have available to us the current and past values of a reference signal, $u[n]$, which is correlated to the desired signal. We also have available to us the past error values of our estimate, $e[n-1] = d[n-1] - \hat{d}[n-1]$. The LMS estimator uses these quantities in the filter structure given in Figure 1 to produce an estimate of the desired signal, $\hat{d}_{lms}[n]$.

The estimate is produced by passing the reference data through an L-tap FIR filter, where the filter weights are updated through the LMS weight update equation,

$$\vec{\mathbf{w}}[n] = \vec{\mathbf{w}}[n-1] + \mu e^*[n-1]\vec{\mathbf{u}}_L[n-1] ,$$

and where

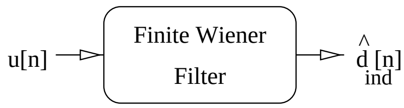

Figure 2: Independence Assumptions Bound

$$\vec{\mathbf{w}}[n] = \left[\begin{array}{c} w_0[n] \ldots w_{L-1}[n] \end{array} \right]^T$$

and

$$\vec{\mathbf{u}}_L[n] = \left[\begin{array}{c} u[n] \ldots u[n-L+1] \end{array} \right]^T \ .$$

Assuming that the initial weight vector at time $n = -\infty$ is the all-zero vector, we can write the LMS estimator as a nonlinear function of the semi-infinite set of reference data, as well as of past values of the desired signal:

$$\begin{aligned} \hat{d}_{lms}[n] &= \mu \sum_{i=-\infty}^{n-1} e[i]\vec{\mathbf{u}}_L^H[i]\vec{\mathbf{u}}_L[n] \quad (1) \\ &= f(u[n], \cdots, u[-\infty], d[n-1], \cdots, d[-\infty]) \ . \end{aligned}$$

A direct analysis of this estimator's performance does not appear to be feasible. The traditional approach to reduce the complexity of the analysis has been to restrict the statistics of input signals through a set of four assumptions, collectively known as the "independence assumptions."

3. INDEPENDENCE ASSUMPTIONS

The performance of any estimator can be bounded by that of the optimal estimator. The optimal MSE estimator is given by the mean of the desired signal, conditioned on the knowledge of all information available to the estimator [1]. Examining the equation for the LMS estimator (1), the optimal estimator is

$$\hat{d}_{opt}[n] = E\left\{d[n]|u[n], \cdots, u[-\infty], d[n-1], \cdots, d[-\infty]\right\} . \quad (2)$$

Actually solving for this estimator requires knowing the statistics of the signals $u[n]$ and $d[n]$. Under the "independence assumptions," the statistics of the signals are restricted through the following four assumptions [3]:

1) tap input vectors $\vec{\mathbf{u}}_L[n], \cdots, \vec{\mathbf{u}}_L[-\infty]$ are independent of each other;

2) $\vec{\mathbf{u}}_L[n]$ is independent of $d[n-1], \cdots, d[-\infty]$;

3) $d[n]$ is dependent on $\vec{\mathbf{u}}_L[n]$, but independent of $d[n-1], \cdots, d[-\infty]$;

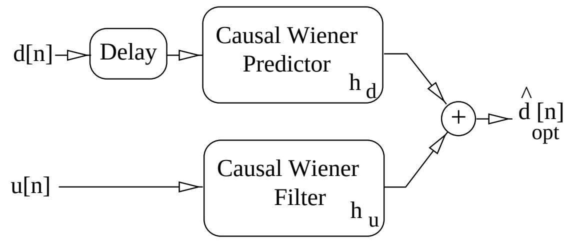

Figure 3: Optimal Estimator

4) $\vec{\mathbf{u}}_L[n]$ and $d[n]$ are mutually Gaussian.

These assumptions simplify the conditional mean, allowing for the optimal estimator to be obtained. Using assumptions 1,2,and 3 reduces the conditional mean to

$$\hat{d}_{ind}[n] = E\left\{d[n]|u[n], \cdots, u[n-L+1]\right\} \ .$$

Condition 4 requires the optimal estimator to be linear, and is then given by

$$\hat{d}_{ind}[n] = \sum_{i=0}^{L-1} w^*_{ind}[i]u[n-i] \ .$$

This is recognized as the finite Wiener filter operating on the L reference values $u[n], \cdots, u[n-L+1]$. Thus, the MSE performance of the LMS estimator, under these assumptions, can be bounded by the MSE of the finite Wiener filter (Figure 2), where the filter weights are given in terms of the autocorrelation matrix of the reference signal and the cross correlation between the reference vector and desired signal.

4. OPTIMAL ESTIMATOR

The "independence assumptions" are very restrictive, exempting most signals found in communications systems. All the cases where the LMS filter outperformed the finite Wiener filter used signals which necessarily violate the "independence assumptions." Since our primary concern is bounding, not modeling, the performance of the LMS filter, we will now derive a bound without using the "independence assumptions."

The optimal estimator is given in (2). Using only an expansion of independence assumption 4, the mutually Gaussian assumption, to include the entirety of both processes, the optimal estimator is a linear estimate and is given by

$$\hat{d}_{opt}[n] = \sum_{i=-\infty}^{n} a_i u[i] + \sum_{i=-\infty}^{n-1} b_i d[i] \ . \quad (3)$$

Note that (3) is a function of all the past reference data, and of all past samples of the desired signal. This equation can be rewritten as the output to the system shown in Figure 3,

$$\hat{d}_{opt}[n] = \sum_{i=-\infty}^{n} h_u[n-i]u[i] + \sum_{i=-\infty}^{n-1} h_d[n-1-i]d[i] \,, \quad (4)$$

where the impulse responses of the causal linear filter and causal linear predictor are given as $h_u[n]$ and $h_d[n]$, respectively. The form of the optimal estimator is known; solving for it requires finding the causal filters, $h_d[n]$ and $h_u[n]$, which result in the minimum MSE. This involves solving the multidimensional Wiener filtering/prediction problem for the given reference and desired signals. The MSE of the optimal estimator bounds the LMS estimator's performance.

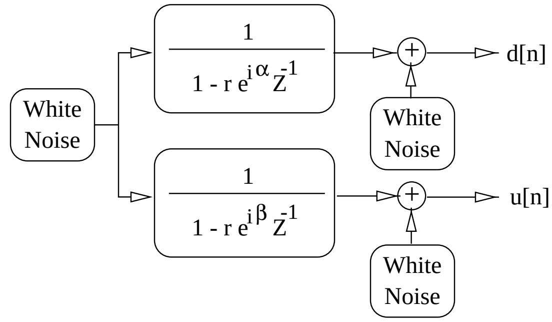

Figure 4: Simulation System

5. RESULTS

To demonstrate this bound for scenarios where the LMS estimator outperforms the finite Wiener filter, we first present a system for which the optimal estimator can be solved, then we produce reference and desired signals which when operated on by an LMS estimator result in the estimator's performance exceeding that of the finite Wiener filter.

The system given in Figure 4 consists of stable autoregressive (AR) processes. The desired and reference signals are first order AR processes generated from the same white Gaussian noise source with an independent Gaussian noise component added to each signal. The optimal estimator (4) for this system can be obtained; the solution involves solving the multidimensional Wiener filtering/prediction problem as presented by Wong [8]. The matrix spectral factorization which this solution requires was performed using a method presented by Davis in [2].

For the LMS adaptive filter to outperform the finite Wiener filter, it is necessary to generate scenarios where the optimal estimator, and thus the LMS filter, have significant contributions from data not available to the finite Wiener filter. This can be achieved by restricting the poles of the AR processes to be close to the unit circle.

The MSE performance of the LMS estimator was evaluated through Monte Carlo simulations. Instead of the standard LMS weight update algorithm, the normalized least mean square (NLMS) algorithm was used [3]. The finite Wiener filter and optimal estimator performances were evaluated numerically. Plots of the MSE performance of the LMS estimator, optimal estimator, and finite Wiener filter are shown under different combinations of the pole locations.

Figure 5 is a graph of the MSE as a function of the adaptation constant μ. For this simulation, the AR pole locations were held constant with radii $r = 0.99$ and an angle separation, $\alpha - \beta$, of 3.6 degrees. From the plot, notice that the LMS estimator does indeed outperform the finite Wiener filter. Also, note that the LMS estimator performance comes close to achieving that of the optimal estimator as μ increases to an optimal value. Having a decreasing MSE as μ increases is contrary to the conventional wisdom that has resulted from the "independence assumptions" model. This can be explained by noting that μ regulates the amount of error fed back into the LMS estimator; by increasing the amount of error being fed back, one is increasing the contributions of the data from the past values of desired and reference data, which are unavailable to the finite Wiener filter.

Figure 6 is a graph of MSE as a function of the pole radii r, where for each value of r the LMS estimator MSE values were taken over a range of adaptation constants; the one which resulted in the smallest MSE was used. The pole angle separation, $\alpha - \beta$, was held constant at 3.6 degrees. The powers of the AR processes were held constant for the various radii by varying the generator noise process's power. The performance of the optimal estimator is almost invariant to the pole locations, and is instead dependent on the signal-to-noise ratios in the reference and desired channels. Also, note that it is not until the poles begin to approach the unit circle that the optimal estimator and the finite Wiener filter's performance begin to diverge. The LMS estimator's performance was dependent on the pole locations, but not to the same degree as the finite Wiener filter; its MSE rose more slowly than that of the finite Wiener filter as the poles approached the unit circle.

Figure 7 is a graph of MSE as a function of the pole angle difference, and again the LMS estimator's performance was evaluated over a range of adaptations constants, the one with the lowest MSE being used. The pole radii were held constant at $r = 0.99$. The smaller the angle separation, the greater cross-correlation between the reference and the desired process. This, along with the long autocorrelation sequences resulting from the poles' proximity to the unit circle, allows the LMS estimator to take advantage of correlated data outside of the range of the finite Wiener filter, resulting in lower MSE for the LMS estimator.

6. CONCLUSION

In conclusion, we bounded the performance of the LMS estimator without using the “independence assumptions.” Cases where the LMS estimator outperforms the finite Wiener filter were then generated, and the performance of the LMS estimator was compared to the bound. The LMS estimator was found to greatly outperform the finite Wiener filter in these cases, and, in some cases, resembled that of the optimal estimator. This behavior can be attributed to the fact that the LMS estimator uses information not available to the finite Wiener filter. This data includes not only all values of the reference data, but also all past values of the desired signal. Finally, while the effect demonstrated in this paper occurs only under severe violation of the “independence assumptions,” it does have applications in both interference suppression and noise cancellation. Both North [4] and Reuter [5] have demonstrated the usefulness of this property when the LMS estimator is used in adaptive equalizers to suppress a narrow-band interferer. As for noise cancellation, the simulations presented in this paper could be easily reworked to represent the LMS estimator when it is used to cancel a Doppler shifted narrow-band signal.

7. REFERENCES

[1] P. J. Brockwell and R. A. Davis, *Time Series: Theory and Methods*, Springer, New York, NY, 2nd edition, 1991.

[2] M. C. Davis, “Factoring the spectral matrix,” *IEEE Transactions on Automatic Control*, vol. 8, no. 4, pp. 296–305, October 1963.

[3] S. Haykin, *Adaptive Filter Theory*, Prentice Hall, Englewood Cliffs, NJ, 3rd edition, 1996.

[4] R. North, R. Axford, and J. Zeidler, “The performance of adaptive equalization for digital communications systems corrupted by interference,” in *27th ASILOMAR Conference on Signals, Systems and Computers*, 1993, pp. 1548–1553.

[5] M. Reuter and J. Zeidler, “Non-Wiener effects in LMS-implemented adaptive equalizers,” in *Proceedings of ICASSP-97*, Munich, Germany, April 1997.

[6] B. Widrow, J. Glover, J. McCool, J. Kaunitz, C. Williams, R. Hearn, J. Zeidler, E. Dong Jr., and R. Goodin, “Adaptive noise canceling: principles and applications,” *Proceedings of the IEEE*, vol. 63, no. 12, pp. 1692–1716, December 1975.

[7] B. Widrow, J. McCool, M. Larimore, and R. Johnson Jr., “Stationary and non-stationary learning characteristics of the LMS adaptive filter,” *Proceedings of the IEEE*, vol. 64, no. 8, pp. 1151–1162, August 1976.

[8] E. Wong, “On the multidimensional prediction and filtering problem and the factorization of spectral matrices,” *Journal of the Franklin Institute*, vol. 272, no. 2, pp. 87–99, August 1961.

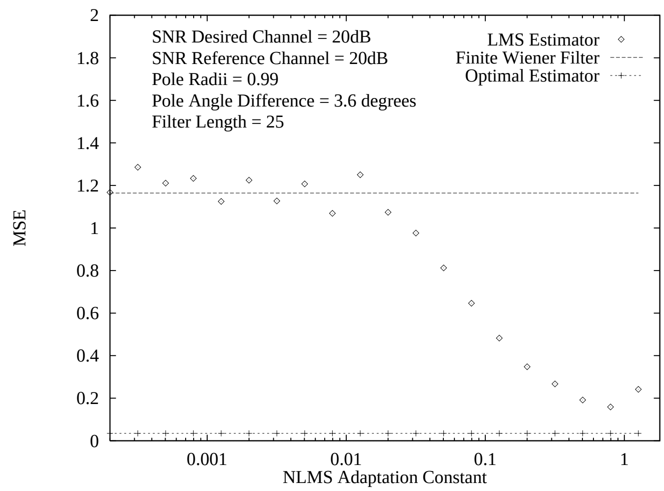

Figure 5: Variation of the NLMS adaptation constant

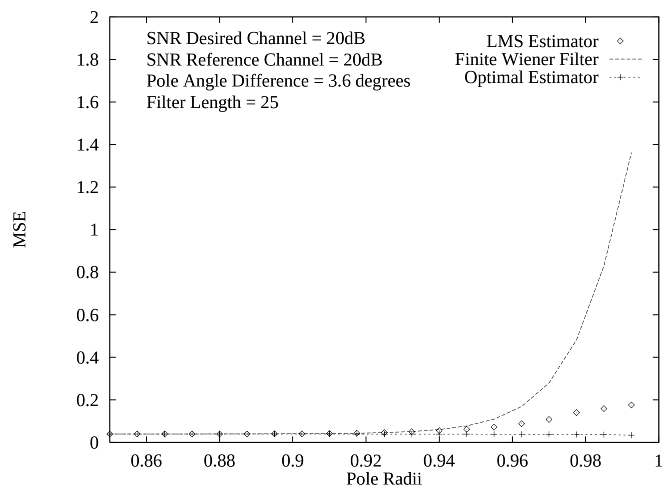

Figure 6: Variation of the pole radii

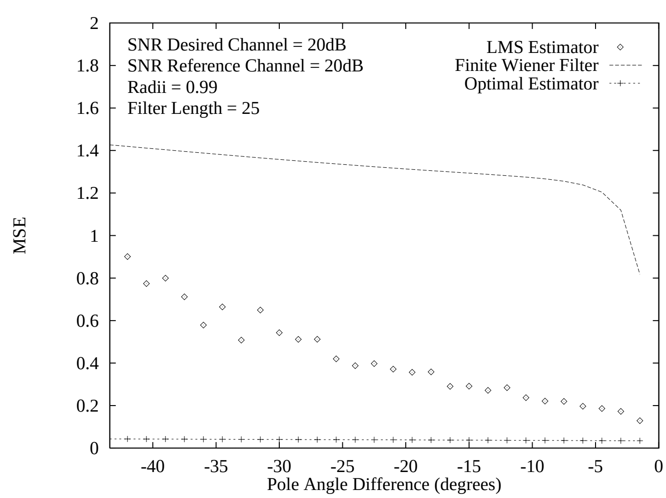

Figure 7: Variation of the pole angle difference