On the Stability and Performance of the Adaptive Line Enhancer

*A. Nehorai**

Department of Electrical Engineering
Stanford University
Stanford, California 94305

*D. Malah***

Department of Electrical Engineering
Technion-Israel Institute of Technology
Haifa, Israel

ABSTRACT

In this presentation the stable and optimal operation of the adaptive line enhancer (ALE) is considered. By improving the estimate of the steady-state mean square error (MSE), a tighter stability constraint is obtained, as well as more accurate expressions for the SNR gain attained by the ALE when filtering sinusoidal signals in white noise. Since the LMS algorithm used for adapting the ALE weights aims at minimizing the MSE, and not at maximizing the output SNR, the proper choice of the algorithm's parameters for maximizing the SNR gain is considered. In particular, it is shown that for a given step-size parameter μ (which satisfies the stability constraint) there exists an optimal number of weights which maximizes the SNR gain. Computer simulations verify the analytical results.

I. Introduction

The adaptive line enhancer (ALE) shown in Fig. 1, which uses the Widrow-Hoff LMS algorithm [1,2] to update the adaptive filter weights, is by now well known [2-5]. It provides efficient means for filtering narrow-band or periodic signals from wide-band signals or noise and is particularly useful for filtering nonstationary signals. The ALE is actually a special case of the more general adaptive noise canceller [2] in which the reference input x_k is a delayed version of the primary input d_k as shown in Fig. 1. By properly choosing the delay duration, the noise components at the two inputs become decorrelated, whereas the narrow-band or periodic signal components remain correlated. Since the weights of the adaptive filter are adapted to minimize the mean square error (MSE), the output of the transversal filter, y_k, is attempting to track the narrow-band components of the input signal. The transversal filter implements therefore a bandpass filter (for a single narrow-band component) or a set of bandpass filters (for multiple narrow-band components). The number of weights, L, determines the bandwidth of each bandpass filter, and hence the improvement or gain in the signal to noise ratio (SNR). However, since the weights are adapted according to an estimate of the MSE gradient, through the LMS algorithm, the noise in the weight-vector $\mathbf{W}$ reduces the performance of the ALE. It is the purpose of this paper to present a better estimate for the steady-state MSE which enables the derivation of more accurate expression for the SNR gain achieved by the ALE, as well as a more accurate stability constraint. Although the improved steady-state MSE estimate and stability constraint coincide with the results obtained by Griffiths [6], they are found here in an extremely simple manner and are applied for evaluating the ALE performance.

Since the LMS algorithm attempts to minimize the MSE, whereas the performance of the ALE is usually measured by the output SNR, we address ourselves also to the problem of choosing the adaptive filter parameters for maximizing the performance of the ALE. Some computer simulations which verify the analytical results are included. Additional details and simulation results can be found in [7].

II. Improved Estimates for the Steady-State MSE and Stability Constraint

The adaptive process according to the LMS algorithm is described by [1,2]

$$\mathbf{W}_{k+1} = \mathbf{W}_k + 2\mu e_k \mathbf{X}_k \tag{1}$$

where $\mathbf{W}_k$ is the weight vector at instant k, $\mathbf{X}_k$ is the reference input vector given by $\mathbf{X}_k = [x_k, x_{k-1}, \cdots x_{k-L+1}]^T$, e_k the error output, and μ is a constant parameter which controls the convergence properties of the algorithm. For sufficiently small μ, the algorithm can be shown to converge in the mean to the Wiener solution [12]

$$\mathbf{W}^* = \mathbf{R}_{xx}^{-1}\mathbf{P} \tag{2}$$

where $\mathbf{R}_{xx}$ is the autocorrelation matrix of the vector $\mathbf{X}_k$ and $\mathbf{P}$ is the cross-correlation vector between he primary input d_k and $\mathbf{X}_k$ (the reference input vector). The usual stability constraint on μ is given by [1,2]

$$0 < \mu < 1/\lambda\max \tag{3}$$

where $\lambda\max$ is the largest eigenvalue of $\mathbf{R}_{xx}$. It will be shown in the sequel that the actual stability range is narrower.

Since the LMS algorithm uses an estimate of the MSE gradient for adapting the weights, the actual instantaneous values of $\mathbf{W}_k$ fluctuate (after convergence) about their mean value $E\{\mathbf{W}_k\} = \mathbf{W}^*$ ("weight-noise") causing a degradation in the performance of the adaptive filter. Assuming that the weights have converged (in the mean) let

$$\mathbf{W}_k = \mathbf{W}^* + \mathbf{V}_k. \tag{4}$$

Then, the output from the transversal filter, y_k, can be described as the sum of two terms

$$y_k = \mathbf{W}_k^T\mathbf{X}_k = \mathbf{W}^{*T}\mathbf{X}_k + \mathbf{V}_k^T\mathbf{X}_k = y_k^* + y_k^v, \tag{5}$$

where y_k^* is the output expected from the optimal Wiener filter and y_k^v is a noise component added due to the weights fluctuations. With the assumption of no correlation between y_k^* and y_k^v (equivalent to the common assumption of no correlation between $\mathbf{X}_k$ and $\mathbf{W}_k$ [1,2]),

* This work is part of an M.Sc. thesis completed at the Department of Electrical Engineering, Technion-Israel Institute of Technology, Haifa, Israel.

** Presently on sabbatical leave at the Acoustic Research Department, Bell Laboratories, Murray Hill, New Jersey 07974.

$$E\{y_k^2\} = E\{(y_k^*)^2\} + E\{(y_k')^2\}. \quad (6)$$

Using the results derived in [2] for the covariance of $\mathbf{V}_k$,

$$E\{(y_k')^2\} = \mu \, tr\mathbf{R}_{xx}\xi_{\min} \quad (7)$$

where $\xi_{\min}$ is the minimum MSE achieved by the Wiener solution. Thus, using (7) the steady-state MSE, ξ_{ss}, is given by

$$\xi_{ss}' = \xi_{\min} + E\{(y_k')^2\} = (1+\mu \, tr\mathbf{R}_{xx})\xi_{\min} \quad (8)$$

and hence the misadjustment M, defined as the ratio of the excess MSE to the minimum MSE, is given by

$$M = \mu \, tr\mathbf{R}_{xx} \quad (9)$$

The results given in (8) and (9) were derived earlier [2,8,9] but were found (in computer simulations that we performed) to be adequate only for very small values of μ (as is actually assumed in the derivations given in the above references). In an attempt to extend the above results for larger values of μ, as well as to adequately predict the divergence of the adaptation process (the upper limit in (3) was found to be too high), we replaced $\xi_{\min}$ in (7) by the actual steady-state MSE, $\tilde{\xi}_{ss}$. Thus in place of (8) we obtain,

$$\tilde{\xi}_{ss} = \xi_{\min} + \mu \, tr\mathbf{R}_{xx}\tilde{\xi}_{ss} \quad (10)$$

and hence

$$\tilde{\xi}_{ss} = \xi_{\min}/(1-\mu \, tr\mathbf{R}_{xx}) \quad (11)$$

resulting in a misadjustment of

$$\tilde{M} = \mu \, tr\mathbf{R}_{xx}/(1-\mu \, tr\mathbf{R}_{xx}) \quad (12)$$

Clearly, if μ is sufficiently small ($\mu tr\mathbf{R}_{xx} << 1$) the results in (11) and (12) coincide with those in (8) and (9), respectively. However, simulation results, such as those shown in Fig. 2, show that indeed, expressions (10)-(12) are adequate for higher values of μ, even up to divergence - which is predicted from (11) to occur when μ reaches $1/tr\mathbf{R}_{xx}$. Thus, the stability constraint on μ which replaces (3) is given by

$$0 < \mu < 1/tr\mathbf{R}_{xx} \quad (13)$$

It is interesting to note that (13) is usually used as a sufficient condition for stability [1,8] since $tr\mathbf{R}_{xx} \geqslant \lambda\max$ and is usually easier to evaluate. The above shows that (13) is also a *necessary* condition.

The increase in ξ_{ss} beyond the value predicted by the usual expression in (8) can be explained (in view of (10) and (11)) as an increase in the MSE through the feedback structure of the adaptive filter which was not taken fully in account in (8). That is, an increase in the MSE is transferred though the LMS algorithm (1) to an increase in the weight-noise causing an additional increase of the MSE. This process is controled by μ and divergence occurs at a lower value of μ than expected from (3).

Finally, we note that expressions identical to (11) and (13) were found by Griffiths [6[. Here they were found, however, through a simple single step of replacing $\xi_{\min}$ in (7) by ξ_{ss}, and additional insight is gained.

III. ALE Performance Evaluation

The modified results in (11)-(13) will now be specifically applied to the evaluation of the performance of the ALE system shown in Fig. 1, when the input d_k is consisting of sinusoidal signals in white stationary noise.

Let the total power of the input signal be P_x. Then, since the reference input signal x_k is a delayed version of the input signal and the transversal filter has L taps,

$$tr\mathbf{R}_{xx} = LP_x. \quad (14)$$

Assuming an input signal of the form

$$d_k = s_k + n_k = \sum_{m=1}^{N} C_m\cos(\omega_m k+\phi_m) + n_k \quad (15)$$

i.e., N sinusoidal signals with an additive zero-mean white noise sequence n_k, the autocorrelation sequence $r_{xx}(l)$ (which determines $\mathbf{R}_{xx}$) is given by

$$r_{xx}(l) = \sum_{m=1}^{N} \frac{C_m^2}{2}\cos\omega_m l + \sigma_n^2\delta(l) \quad (16)$$

where σ_n^2 is the noise power and $\delta(l)$ is the Kronicker δ-function ($\delta(l)=1$ for $l=0$; $\delta(l)=0$, otherwise). Assuming that L is sufficiently large ($L >> 2\pi/(\omega_r-\omega_p)$ for any $r \neq p$; $r,p\epsilon[1,N]$) the optimal Wiener solution $\mathbf{W}^*$ can be described by the sum [3]

$$\mathbf{W}^* = \sum_{m=1}^{N} \mathbf{W}_m^* \quad (17)$$

where $\mathbf{W}_m^*$ is the Wiener solution for a single sinusoidal signal in white noise (at frequency ω_m) given by [2,3]

$$\mathbf{W}_m^* = \frac{2}{L}a_m^*[\cos\omega_m\Delta,\cos\omega_m(\Delta+1),\ldots,\cos\omega_m(\Delta+L-1)]^T \quad (18)$$

where

$$a_m^* = (\frac{L}{2})\rho_{im}/(1+\rho_{im}L/2) \quad (19)$$

and ρ_{im} is the input SNR for the m-th sinusoidal component, i.e.

$$\rho_{im} = C_m^2/2\sigma_n^2 \quad (20)$$

The corresponding sinusoidal component at the output of the transversal filter is given by

$$s_{ym}(k) = a_m^*(C_m\cos(\omega_{mk}+\phi_m')) \quad (21)$$

The total power of the output signal from the transversal filter, having the ideal weights $\mathbf{W}^*$, is therefore given by

$$E\{(y_k^*)^2\} = \sigma_n^2\frac{2}{L}\sum_{m=1}^{N}(a_m^*)^2 + \sum_{m=1}^{N}(a_m^*C_m)^2/2 \quad (22)$$

The overall output SNR is given therefore by

$$\rho_O^* = \frac{L}{2}\left[\sum_{m=1}^{N}(a_m^*C_m)^2/2\right]/\left[\sigma_n^2\sum_{m=1}^{N}(a_m^*)^2\right]$$
$$= \frac{L}{2}\sum_{m=1}^{N}\rho_{im}(a_m^*)^2/\sum_{m=1}^{N}(a_m^*)^2 \quad (23)$$

where ρ_{im} is defined in (20) and a_m^* is given in (19). The overall input SNR, ρ_i, is given by

$$\rho_i = \sum_{i=1}^{N}\rho_{im}$$

and we define,

$$\Gamma^* \triangleq \rho_0^*/\rho_i = \frac{L}{2}\sum_{m=1}^{N}\rho_{im}(a_m^*)^2/\rho_i\sum_{m=1}^{N}(a_m^*)^2\,. \quad (24)$$

Γ^* is the gain in SNR achieved by the ALE which has the Wiener-solution weights.

For the particular case of equal power sinusoids, $\rho_{im} = \rho_i/N$, $m = 1,2,\ldots,N$,

$$\Gamma^* = L/(2N) \quad (25)$$

The decrease in Γ^* with the increase in N is due to the corresponding larger number of bandpass filters, each passing not only the desired signal but also a band of the noise, thus increasing the overall output noise.

We turn now to the performance of the ALE with the actual weights $\mathbf{W}$ as obtained with the LMS algorithm. From (10), (11) and (12), we conclude that in order to find the actual total output power one has to add to the right hand side of (22) an additional term which is equal to the excess MSE given, by $\tilde{M}\xi_{\min}$. Thus (23) is replaced by

$$\rho_0 = \frac{L}{2}\sum_{m=1}^{N}\rho_{im}(a_m^*)^2/\left[\sum_{m=1}^{N}(a_m^*)^2+\tilde{M}\xi_{\min}\right], \quad (26)$$

where for the evaluation of (26) one should use (12) and (14) for $\tilde{M}$ and the following expression for ξ_{min}

$$\xi_{min} = \sigma_n^2\left[1+\frac{2}{L}\sum_{m=1}^{N}(a_m^*)^2\right] + \sum_{m=1}^{N}(C_m^2/2)(1-a_m^*)^2. \qquad (27)$$

For the particular case of equal-power sinusoids so that $\rho_i = \rho_{in}/N$ and $a^* = a_m^*$, $m = 1,2,...,N$, one obtains

$$\Gamma \triangleq \rho_0/\rho_i =$$

$$(a^*)^2/\left\{\frac{2N}{L}(a^*)^2+[\mu Lp_x/(1-\mu Lp_x)][1+\frac{2N}{L}(a^*)^2+\rho_i(1-a^*)^2]\right\} \qquad (28)$$

With the substitution of (19) for a^* in (28) (using $\rho_i = \rho_{in}/N$)

$$\Gamma = 1/\left\{\frac{2N}{L}+\frac{\mu LP_x}{1-\mu LP_x}[1+\frac{2N}{L}(1+\frac{2}{\rho_i})+\frac{4N^2}{L^2\rho_i}(1+\frac{1}{\rho_i})]\right\} \qquad (29)$$

For the case of a single sinusoidal input, one simply has to substitute $N = 1$ in the above expression. The result (for N=1) coincides with the one obtained in [4] if $\mu LP_x << 1$ and $\rho_i << 1$. In other cases (29) is more accurate. Fig. 3 shows a comparison between the analytical results and computer simulations for $N = 1$. The ripple in the simulation results is apparently due to the effect of coupling between the positive and negative frequency components of the real sinusoid [3]. As the frequency of the input sinusoid is fixed and L varies, the coupling is periodically zeroed resulting in peaks in the performance curve (see also Eqn. (26) in [3]).

According to (13), the stability of the ALE is assured for

$$0 < \mu < 1/LP_x \qquad (30)$$

For a single sinusoid of amplitude C: $LP_x = L(\sigma_n^2+C^2/2)$, where as λmax can be shown [5] to be given by $\lambda max = \sigma_n^2 + (L/2)(C^2/2)$ so that (3) can be significantly different from (12). The validity of (30) is verified in Fig. 2 where clearly divergence occurs when μLP_x approaches 1.

IV. Optimal Choice of ALE Parameters

As mentioned earlier, since the LMS algorithm attempts to minimize the MSE it does not maximize, in general, the output SNR as would be desired for the ALE. This can be seen from

$$\min_{W} E\{e_k^2\} = E\{n_k^2\} + \min_{W} E\{(s_k-s_{y_k})^2+n_{y_k}^2\} \qquad (31)$$

which is clearly not equivalent to maximizing ρ_O, where

$$\rho_O = E\{s_{y_k}^2\}/E\{n_{y_k}^2\}. \qquad (32)$$

It is therefore of importance to properly choose the number of weights L and the step-size parameter μ in order to optimize the performance of the ALE for a given application.

Due to the difficulty in obtaining general expressions for N sinusoidal signals with unequal amplitudes, we will narrow down our discussion to the case of N equal power sinusoids (which includes the case of a single sinusoid). In the previous section we have seen that the Wiener solution results in $\Gamma^* = L/2N$. Hence, from (28), to approach this result μLP_x should be made sufficiently smaller than 1. To obtain a higher value for Γ^*, L should be increased and μ decreased accordingly (the maximum is reached for $L \to \infty$ and $\mu \to 0$). In practice L cannot be increased beyond a certain L_{max} and μ cannot be decreased below a certain $\mu_{min} > 0$. The latter is due to the finite word-length of the digital representation of μ and the arithmatic operations, or due to a given requirement on the adaptation time-constant. Let us consider first the case in which μ is set to $\mu = \mu_0$ (e.g. $\mu_0 = \mu_{min} > 0$). It can be shown that in such a case there exist a finite value of L which maximizes the SNR gain Γ. With the practical assumption that $\mu_0 P_x L << 1$ the optimal value for L is found, by differentiating (29) with respect to L, to be

$$L_{opt} = \left[\frac{2N}{\mu_0 P_x} + \frac{4N}{\rho_i^2} + \frac{4N}{\rho_i}\right]^{1/2} \qquad (33)$$

If also $2\mu_0 P_x << \rho_i^2/(\rho_i+1)$ (which is often the case), (33) is simplified to

$$L_{opt} \cong [2N/(\mu_0 P_x)]^{1/2} \qquad (34)$$

The maximum SNR gain is then given by

$$\Gamma_{max} \cong L_{opt}/(4N) \qquad (35)$$

Note that $\Gamma_{max} = \Gamma^*/2$ for the particular values of $\mu = \mu_0$ and $L = L_{opt}$. Fig. 4 shows the behavior of Γ as a function of L for different values of μ and ρ_i. The existance of an optimal L is clearly seen. This behavior is also seen in Fig. 3. The stability constraint in (30) is satisfied, when L is chosen according to (34), if $L/2 > N$. However, according to our assumption above that $\mu_0 P_x L << 1$, the above results are accurate if $L_{opt}/2 >> N$.

A most important practical consideration is the adaptation time-constant τ_w of the weights, which determines the convergence-time and the ability of the filter to track nonstationary signals. For N equal power sinusoids in white noise, τ_w is given by [5]

$$\tau_w = 1/[2\mu(\sigma_n^2+(L/2N)\sigma_s^2)] \qquad (36)$$

where σ_s^2 is the total power of the N sinusoidal signals and τ_w is measured in terms of adaptation iterations. Thus, if $\tau_w = \tau_0$ is specified, (36) gives the relation between μ and L and once anyone of them is set, the other is determined. An attempt to maximize Γ in (29) subject to the given relation in (36) reveals again that Γ is maximized for the unacceptable values of $\mu \to 0$ and $L \to \infty$. A practical choice appears to be the use of $L = L_{max}$, the maximum possible L, and setting μ to satisfy (36) for the given $\tau_w = \tau_0$ and $L = L_{max}$.

Finally, if the N sinusoidal components do not have equal power, one can still use the above results as guidelines for choosing the ALE parameters. This is based on the facts that the output SNR is mainly determined by the stronger components and that the weaker ones have longer adaptation time-constants. Thus, the given specifications can be modified accordingly and the above results can be then used.

V. Conclusion

By incorporating the effect of the feedback structure, to its full extent, into the analysis of LMS algorithm weight-vector noise and its contribution to the steady state MSE, a better estimate of the misadjustment and a stricter stability range were derived for the adaptive noise-canceller. These results are then applied for studying the performance of the ALE in filtering sinusoidal signals from white noise. The performance measure used is the SNR gain, Γ, achieved by the ALE. Guidelines for the optimal selection of the ALE parameters, namely the number of weights L and the adaptation step-size parameter μ, are given. In particular, it is found that for a given value of μ there exists a finite optimal value of L which maximizes Γ. The value of Γ is then half of the value obtained by the optimal Wiener solution.

This study is a part of a broader study relating to the application of the ALE structure, in the time and frequency domains, for filtering speech signals from noise [7].

References

[1] B. Widrow, "Adaptive filters," in *Aspects of Network and System Theory*, R. Kalman and N. Declaris, Eds., New York: Holt, Rinehart and Winston, 1971, pp. 563-587.

[2] B. Widrow et al., "Adaptive noise cancelling: principles and applications," Proc. IEEE Vol. 63, No. 12, pp. 1692-1716, December 1975.

[3] J. R. Zeidler et al., "Adaptive enhancement of multiple sinusoids in uncorrelated noise," IEEE Trans. Acoust., Speech, Signal Processing, Vol. ASSP-26, NO. 3, pp. 240-254, June 1978.

[4] J. T. Rickard and J. R. Zeidler, "Second-order output statistics of the adaptive line enhancer," IEEE Trans. Acoust., Speech, Signal Processing, Vol. ASSP-27, no. 1, pp. 31-39, February 1979.

[5] J. Treichler, "Transient and convergent behavior of the adaptive line enhancer," IEEE Trans. Acoust., Speech, Signal Processing, Vol. ASSP-27, No. 1, pp. 53-63, February 1979.

[6] L. J. Griffiths, "Rapid measurement of instantaneous frequency," IEEE Trans. Acoust., Speech, Signal Processing, Vol. ASSP-23, No. 2, pp. 209-222, April 1975.

[7] A. Nehorai, "Adaptive filtering of speech signals from noise," M.Sc. thesis (in Hebrew), Technion-IIT, August 1979.

[8] B. Widrow et al., "Stationary and nonstationary learning characteristics of the LMS adaptive filter," Proc. IEEE, Vol. 64, No. 8, pp. 1151-1162, August 1976.

[9] B. Widrow and J. McCool, "A comparison of adaptive algorithms based on the methods of steepest descent and random search," IEEE Trans. Antennas and Propagation, Vol. AP-24, No. 5, pp. 615-637, September 1976.

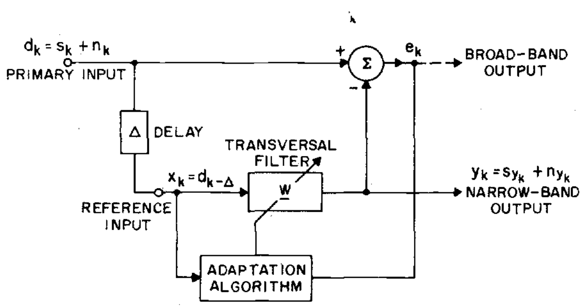

Fig. 1. Block diagram of the ALE.

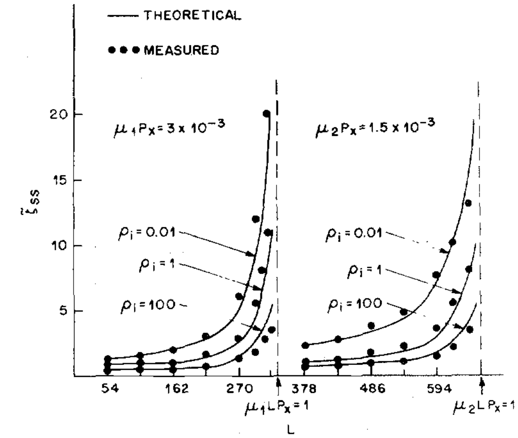

Fig. 2. Comparison of analytical and simulation results for the steady state MSE $\tilde{\xi}_{ss}$ of the ALE with a single sinusoidal input signal in white noise.

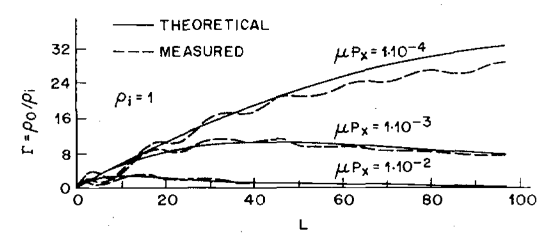

Fig. 3. Theoretical and measured SNR gain Γ of the ALE as a function of the number of weights L.

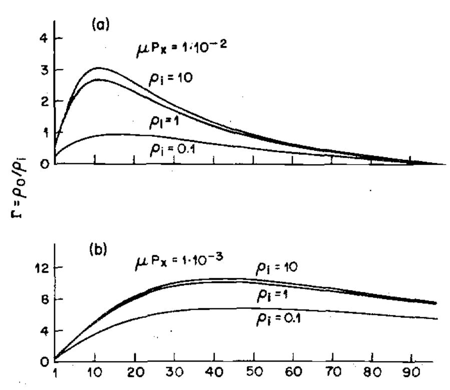

Fig. 4. The dependence of Γ (in (29)) on L for various values of μ and ρ_i.