

Nonlinear Effects in LMS Adaptive Equalizers

Michael Reuter, *Member, IEEE*, and James R. Zeidler, *Fellow, IEEE*

Abstract—An adaptive transversal equalizer based on the least-mean-square (LMS) algorithm, operating in an environment with a temporally correlated interference, can exhibit better steady-state mean-square-error (MSE) performance than the corresponding Wiener filter. This phenomenon is a result of the nonlinear nature of the LMS algorithm and is obscured by traditional analysis approaches that utilize the independence assumption (current filter weight vector assumed to be statistically independent of the current data vector). To analyze this equalizer problem, we use a transfer function approach to develop approximate analytical expressions of the LMS MSE for sinusoidal and autoregressive interference processes. We demonstrate that the degree to which LMS may outperform the corresponding Wiener filter is dependent on system parameters such as signal-to-noise ratio (SNR), signal-to-interference ratio (SIR), equalizer length, and the step-size parameter.

Index Terms—Adaptive equalizers, adaptive filters, least mean square methods, nonlinearities, Wiener filtering.

I. INTRODUCTION

ADAPTIVE transversal equalizers are important components of digital receivers and are primarily used to compensate for the effects of intersymbol interference in bandwidth-constrained communication channels [1]. An additional scenario of some practical interest, particularly in the mobile communication environment, is the operation of the adaptive equalizer in the presence of an interferer [2], [3].

The computationally efficient least-mean-square (LMS) adaptive algorithm [4] is often used in the implementation of the equalizer. Due to the nonlinearity of the LMS algorithm, the optimum performance of the equalizer is often assessed using the Wiener realization of the adaptive filter. The efficacy of this approach is based on the argument that the LMS algorithm will result in greater mean-square-error (MSE) than the corresponding Wiener filter due to gradient noise on the adaptive filter weights. This argument is supported by traditional analysis approaches that invoke the independence assumption in which it is assumed that the current filter weight vector is statistically independent of the current tap data vector [5], [6, pp. 392–399]. Then, the resulting analytical expression of the MSE of the LMS algorithm is greater than the MSE produced by the Wiener filter. The expressions

derived using the independence assumption generally have agreed closely with experimental results for a variety of adaptive filter applications such as the adaptive line enhancer and the adaptive noise canceler [4], [7]–[10] when the LMS step-size parameter has a “small value” [4], [5]. Mazo also has justified the independence assumption for the equalizer application for a small step-size parameter [11].

However, it has recently been reported that an LMS-implemented adaptive equalizer operating with a temporally correlated interferer can produce better probability-of-error performance than the corresponding Wiener filter [12]. Subsequent simulations have revealed the unexpected result that with the proper choice of the step-size parameter, the nonlinear nature of the LMS algorithm can be exploited to generate MSE that is less than the Wiener MSE. This effect occurs if the interferer bandwidth is much less than the bandwidths of the communication signal and additive noise process, resulting in strong coupling between the LMS weight vector and the input reference data. As a result, an analysis of this problem cannot invoke the independence assumption.

Recently, approaches have been proposed to analyze the LMS algorithm without using the independence assumption. Douglas [13] has presented an exact expectation approach that results in a set of linear equations for predicting LMS performance. However, for filter lengths applicable to this equalizer problem, the number of equations becomes computationally burdensome. Haykin [6, App. I] and Butterweck [14] have proposed a method that uses a power series expansion of the LMS weights. The mathematics are made tractable by making a small step-size parameter assumption. This approach reveals structure in the statistical characteristics of the LMS algorithm that is not observed when the independence assumption is used. However, the simplifying assumptions inherent in this approach are invalid for values of the step-size parameter that produces the most pronounced nonlinear characteristics of the LMS algorithm.

To analyze this behavior, we utilize the transfer function approach first presented by Glover [15] for adaptive noise canceling of sinusoidal interferences and later generalized by Clarkson and White [16] to include deterministic interferences of arbitrary periodic nature and interferences that are stochastic. We present an analysis approach that generates an approximate expression of the steady-state MSE for the LMS algorithm and demonstrate how it can be modified to include the normalized LMS algorithm (NLMS). We specifically analyze equalizer performance for interference that is sinusoidal and an autoregressive process of order one [AR(1)]. For the sinusoidal interference scenario, our results illustrate the relationship between the performance improvement of the

Manuscript received June 7, 1996; revised July 16, 1998. This work was supported by the Independent Research Program of the Space and Naval Warfare Systems Center, San Diego, CA, and by a Grant of High Performance Computing Time from the DoD HPC Center. The associate editor coordinating the review of this paper and approving it for publication was Prof. Tyseer Aboulnasr.

The authors are with the Space and Naval Warfare Systems Center, San Diego, CA 92152 USA. They are also with the Department of Electrical and Computer Engineering, University of California at San Diego, La Jolla, CA 92093 USA.

Publisher Item Identifier S 1053-587X(99)03644-2.

LMS algorithm and system parameters such as signal-to-noise ratio (SNR), equalizer length, and the step-size parameter. We use this expression of MSE to determine the optimum step-size parameter that maximizes the performance improvement. We analyze the AR(1) interference scenario in a similar fashion, except in this case, the equalizer is implemented with the NLMS algorithm.

In Section II, we describe the equalizer problem and introduce notation. In Section III, we present the general transfer function approach for approximating the MSE of the LMS and NLMS algorithms. Section IV contains the specific solutions for the sinusoidal interference and the AR(1) interference. We also include some numerical examples. We make some concluding remarks in Section V.

II. EQUALIZER PROBLEM

Lower-case bold letters represent vectors, upper-case bold letters represent matrices, and all other data quantities are scalars. In general, the data are complex.

Fig. 1 represents the baseband adaptive equalizer structure to be analyzed. For the sake of brevity, we present only the analysis of a finite, symmetric, two-sided equalizer, even though the nonlinear effects occur in one-sided equalizers as well. Vector quantities such as the reference data vector $\mathbf{u}(k)$ are represented as

$$\mathbf{u}^T(k) = (u(k+N) \quad u(k+N-1) \quad \dots \quad u(k) \quad \dots \quad u(k-N+1) \quad u(k-N)) \quad (1)$$

where

- N number of precursor and postcursor taps;
- k time index;
- T transpose.

The total number of taps is given by $L = 2N + 1$. $\mathbf{u}(k)$ can be decomposed into a sum of three statistically independent components as

$$\mathbf{u}(k) = \mathbf{s}(k) + \mathbf{x}(k) + \mathbf{n}(k) \quad (2)$$

where

- $\mathbf{s}(k)$ communication signal vector;
- $\mathbf{x}(k)$ interference vector;
- $\mathbf{n}(k)$ noise vector.

The scalar $u(k) = s(k) + x(k) + n(k)$ is the sample of the reference process residing at the center tap of the transversal filter at time index k . All processes are modeled as wide sense stationary, with zero mean. Because the bandwidth of the interference process $x(k)$ is assumed to be more narrow than the bandwidth of the noise process, the interference and noise terms are separated in (2).

The communication signal $s(k)$ and the noise process $n(k)$ are modeled as white (samples are mutually independent). In this scenario, the equalizer is not being used to compensate for channel-induced distortion of the communication signal but is only being used to mitigate the effects of the additive narrowband interference [2], [3]. We make this assumption for the purpose of simplifying the presentation, because the narrowband interference is the cause of the nonlinear phenomenon

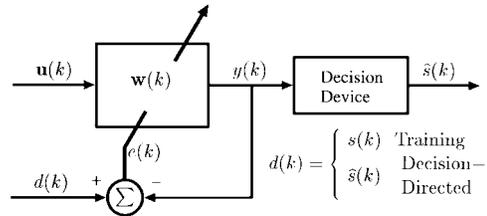


Fig. 1. Adaptive equalizer structure.

of the LMS algorithm. This behavior also is observed when $s(k)$ has a nonwhite spectrum. Moreover, including the effects of a dispersive communication channel into the analysis is possible.

The output of the adaptive filter $y(k)$ is given by the inner product of the filter weights and the data vector as

$$y(k) = \mathbf{w}^H(k)\mathbf{u}(k) \quad (3)$$

where H denotes Hermitian transpose. $y(k)$ is sent through a decision device to estimate $s(k)$, which is the communication symbol currently at the center tap.

During the convergence phase of the adaptive algorithm, the equalizer is in the training mode, in which the desired or primary input $d(k)$ is the error-free training sequence $s(k)$. During the communication phase, the equalizer is in the decision-directed mode, in which the output of the decision device $\hat{s}(k)$ is used as $d(k)$.

The weights of the equalizer are adapted by the complex LMS algorithm, which uses the error sequence $e(k)$ to adjust the weights as [4]

$$\mathbf{w}(k+1) = \mathbf{w}(k) + \mu\mathbf{u}(k)e^*(k) \quad (4)$$

where μ is the step-size parameter, $e(k) = d(k) - y(k)$, and $*$ denotes complex conjugation.

Defining the $L \times L$ correlation matrix and the $L \times 1$ cross-correlation vector as

$$\mathbf{R} = E[\mathbf{u}(k)\mathbf{u}^H(k)] \quad \text{and} \quad \mathbf{p} = E[\mathbf{u}(k)d^*(k)] \quad (5)$$

the Wiener filter weights and associated MSE are given by [6, ch. 5]

$$\mathbf{w}_w = \mathbf{R}^{-1}\mathbf{p} \quad \text{and} \quad J_w = \sigma_s^2 - \mathbf{p}^H\mathbf{w}_w \quad (6)$$

where $J_w = E[|d(k) - \mathbf{w}_w^H\mathbf{u}(k)|^2]$, and σ_s^2 is the communication signal power. Using the statistical assumptions on the components of $\mathbf{u}(k)$ and assuming that the desired signal $d(k)$ is equal to the communication signal $s(k)$, i.e., there are no decision errors, \mathbf{p} can be shown to be [3]

$$\mathbf{p}^T = \sigma_s^2(0 \quad \dots \quad 0 \quad 1 \quad 0 \quad \dots \quad 0) \quad (7)$$

where the $(N+1)$ -element is the only nonzero term. Equation (7) is strictly true only during the training phase of the equalizer operation. However, we also assume that there are no decision errors when the equalizer is in the decision-directed mode. This simplifying assumption is commonly made in equalizer performance studies [17, pp. 593–598] and, generally, is required because little is known about the joint distribution of the feedback errors [18]. This assumption

usually is tested against Monte Carlo simulations to determine when decision feedback errors create serious discrepancy with theory.

The equalizer problem consists of determining the minimum mean-square estimate of the communication symbol $s(k)$. If $s(k)$ is estimated only by a function of the reference data vector $\mathbf{u}(k)$, the general optimum estimate which minimizes MSE is given by [19, pp. 412–413]

$$C_o(\mathbf{u}(k)) = E[s(k)|\mathbf{u}(k)] \quad (8)$$

and the optimum linear estimate is given by the Wiener filter output

$$C_w(\mathbf{u}(k)) = \mathbf{w}_w^H \mathbf{u}(k). \quad (9)$$

If the data are all Gaussian distributed with zero mean (including the communication signal), then the Wiener filter is the optimum MSE estimator.

However, this fact does not preclude the LMS algorithm from generating an estimate with less MSE, even under the physically unrealistic assumption that the communication signal is Gaussian. The LMS estimate of $s(k)$ not only is a function of the current reference vector $\mathbf{u}(k)$, but because of the recursive nature of the update equation of (4), it is a function of past reference data $\mathbf{u}(k-1)$, $\mathbf{u}(k-2) \dots$ as well. More importantly, it uses previously detected samples of the communication signal $d(k-1)$, $d(k-2) \dots$, which enter in through the error term. The Wiener estimate is only a function of the cross-correlation between the communication signal and the reference data. The LMS estimator can be written abstractly as

$$y(k) = C_{\text{lms}} \begin{pmatrix} \mathbf{u}(k) & \mathbf{u}(k-1) & \mathbf{u}(k-2) & \dots \\ d(k-1) & d(k-2) & \dots \end{pmatrix}. \quad (10)$$

Therefore, the LMS estimate of $s(k)$ is a function of much more information than that used by the Wiener filter of (9).

In fact, this information commonly is used in a decision-feedback equalizer structure consisting of a feedforward transversal filter that processes the reference data with a decision-feedback filter that processes the previously detected communication symbols [17, pp. 593–598]. The purpose of the feedback filter is to remove any residual intersymbol interference caused by previously transmitted symbols. However, the transversal equalizer of Fig. 1 implemented with the LMS algorithm is not equivalent to the decision-feedback equalizer because they process the information differently.

As another method to compare with the LMS estimator, we consider a decision-feedback equalizer consisting of a feedforward filter with N precursor and postcursor taps as in the transversal equalizer of Fig. 1, with an additional N -tap feedback filter in which the previously-detected symbols are explicitly incorporated. This estimator can be written as

$$C_{df}(\mathbf{u}(k); \mathbf{d}(k)) = \mathbf{w}_{df}^H \mathbf{u}(k) + \mathbf{b}^H \mathbf{d}(k) \quad (11)$$

where

$$\mathbf{d}^T(k) = (d(k-1) \quad d(k-2) \quad \dots \quad d(k-N)). \quad (12)$$

\mathbf{w}_{df} represents the feedforward taps and \mathbf{b} the feedback taps. When the feedforward and feedback weights are jointly optimized according to the minimum MSE criterion and there are no decision errors, the Wiener filter weights are given by [1]

$$\mathbf{w}_{df} = (\mathbf{R} - \mathbf{Q}^H \mathbf{Q})^{-1} \mathbf{p} \quad \text{and} \quad \mathbf{b} = -\mathbf{Q} \mathbf{w}_{df} \quad (13)$$

where

$$\mathbf{Q} = E[\mathbf{d}(k) \mathbf{u}^H(k)]. \quad (14)$$

The associated MSE defined by $J_{df} = E[|d(k) - \mathbf{w}_{df}^H \mathbf{u}(k) - \mathbf{b}^H \mathbf{d}(k)|^2]$ is

$$J_{df} = \sigma_s^2 - \mathbf{p}^H \mathbf{w}_{df}. \quad (15)$$

Even though the LMS estimator of (10) and the Wiener estimator with decision feedback of (11) strictly do not use the same information, it is instructive to compare their performance to determine how effectively the LMS algorithm uses previously detected symbols to enhance performance over the Wiener estimator of (9).

III. TRANSFER FUNCTION APPROACH

We begin by using the standard method of decomposing the LMS filter weights into a sum of the time-invariant Wiener filter and a time-varying misadjustment component as [7]

$$\mathbf{w}(k) = \mathbf{w}_w + \mathbf{w}_{mis}(k). \quad (16)$$

We start filtering at time index $k = 0$ and assume that $\mathbf{w}_{mis}(0) = \mathbf{0}$. This is equivalent to initializing the LMS algorithm with the Wiener filter. Using (3), the output $y(k)$ can be decomposed as

$$y(k) = y_w(k) + y_{mis}(k) \quad (17)$$

where $y_w(k)$ is the output of the Wiener filter, and $y_{mis}(k)$ is the output of the misadjustment filter. Using (4), we get the recursive equation for the misadjustment filter

$$\begin{aligned} \mathbf{w}_{mis}(k+1) &= \mathbf{w}_{mis}(k) + \mu \mathbf{u}(k) e^*(k) \\ &= \mu \sum_{j=0}^k e^*(j) \mathbf{u}(j). \end{aligned} \quad (18)$$

Then, the output process $y(k)$ can be written as

$$y(k) = \mathbf{w}_w^H \mathbf{u}(k) + \mu \sum_{j=0}^{k-1} e(j) \mathbf{u}^H(j) \mathbf{u}(k) \quad (19)$$

while the error process $e(k) = d(k) - y(k)$ satisfies the recursive difference equation

$$e(k) + \mu \sum_{j=0}^{k-1} e(j) \mathbf{u}^H(j) \mathbf{u}(k) = d(k) - \mathbf{w}_w^H \mathbf{u}(k). \quad (20)$$

Equation (20) is a k th-order recursive difference equation. Because the coefficients are stochastic, this equation is difficult to solve analytically. However, for wide sense stationary processes whose second-order moments can be estimated with

time averages, and for large enough filter length ($L \gg 1$), Clarkson and White [16] propose using the approximation

$$\mathbf{u}^H(j)\mathbf{u}(k) \approx Lr_u(k-j) \quad (21)$$

where $r_u(m)$ is the autocorrelation function of the reference process $u(k)$. Equation (20) then is approximated by a standard difference equation with constant coefficients as

$$e(k) + \mu L \sum_{j=0}^{k-1} r_u(k-j)e(j) = e_w(k) \quad (22)$$

where

$$e_w(k) = d(k) - \mathbf{w}_w^H \mathbf{u}(k) \quad (23)$$

is the residual Wiener error process with power J_w .

We then interpret the steady-state ($k \rightarrow \infty$) LMS error $e(k)$ as the output of a time invariant linear system with transfer function $H_E(z)$ given by [16]

$$H_E(z) = \frac{1}{1 + \mu LR(z)} \quad (24)$$

with

$$R(z) = \sum_{m=1}^{\infty} r_u(m)z^{-m} \quad (25)$$

driven by the wide sense stationary Wiener error process $e_w(k)$. The discrete power spectrum of the error process $e(k)$ is

$$S_E(z) = H_E(z)H_E^*(1/z^*)S_{E_w}(z). \quad (26)$$

The discrete power spectrum of the Wiener error process is derived using (2) and (23) and is given by

$$S_{E_w}(z) = (1 - W_w^*(z^*))(1 - W_w(1/z))S_s(z) + W_w^*(z^*)W_w(1/z)(S_x(z) + S_n(z)) \quad (27)$$

where $S_s(z)$, $S_x(z)$, and $S_n(z)$ are the discrete power spectra of the communication signal, the interference process, and the noise process, respectively, and $W_w(z)$ is the transfer function of the Wiener filter. Again, we assume that even if the equalizer is in the decision-directed mode, there are no decision errors. Then, the approximation of the steady-state MSE of the LMS algorithm is the power of the error process $e(k)$ given by

$$J_{\text{lms}} = \frac{1}{2\pi i} \oint_{|z|=1} S_E(z)z^{-1} dz. \quad (28)$$

The transfer function approach allows for the communication signal $s(k)$ and the noise process $n(k)$ to have arbitrary spectral properties. As a consequence, this approach accommodates analysis of the more general problem in which the dispersive nature of the communication channel induces temporal correlation in the signal $s(k)$. Therefore, it is possible for channel equalization to be incorporated in a straightforward manner. The Wiener filter would have to be calculated accordingly. In particular, the cross correlation vector of (7) would no longer be valid in general. However, because we are restricting the analysis to white processes, in (27), we replace

$S_s(z)$ and $S_n(z)$ with the respective signal and noise powers σ_s^2 and σ_n^2 .

This transfer function approach can also be applied to the NLMS algorithm [6, pp. 432–437]. We use the NLMS version in which the weights are adapted as

$$\hat{\mathbf{w}}(k+1) = \hat{\mathbf{w}}(k) + \frac{\hat{\mu}}{\|\mathbf{u}(k)\|^2} \mathbf{u}(k)e^*(k). \quad (29)$$

Using (21), we approximate the square of the norm of the reference data vector as

$$\|\mathbf{u}(k)\|^2 \approx L(\sigma_s^2 + \sigma_x^2 + \sigma_n^2) \quad (30)$$

where σ_x^2 is the interference power. Then, we use (26)–(28) to estimate the MSE of the NLMS algorithm J_{nlms} by replacing $H_E(z)$ with $\hat{H}_E(z)$ given by

$$\hat{H}_E(z) = \frac{1}{1 + \hat{\mu}R(z)/(\sigma_s^2 + \sigma_x^2 + \sigma_n^2)}. \quad (31)$$

The decomposition of the LMS filter weights of (16) into the Wiener and misadjustment components is an important aspect of the transfer function approach considered in this paper and is a fundamental difference with the method of [16]. Conventional analysis of the LMS algorithm interprets the misadjustment filter $\mathbf{w}_{\text{mis}}(k)$ as the cause of *excess* MSE. In contrast, for the equalizer scenario considered in this paper, the contribution of $\mathbf{w}_{\text{mis}}(k)$ to the composite LMS filter of (16) can result in MSE, which is less than that of the Wiener filter alone. The significance of $\mathbf{w}_{\text{mis}}(k)$ to this effect was first demonstrated in [12]. This is a nonlinear phenomenon because the misadjustment weights are dependent on the reference as well as the desired input data. In addition, by analytically initializing the LMS algorithm with the Wiener weights, we mitigate transient effects associated with initial convergence, thereby improving the accuracy of the difference equation of (22) in approximating the steady-state behavior of the exact equation in (20). Although some transient effects remain, especially if the LMS time-invariant steady-state weight vector component is not well represented by the Wiener filter, the decomposition of (16) has resulted in theoretical MSE calculations that adequately model actual LMS performance.

IV. EXAMPLES

We apply the results of the previous section to two interference scenarios. First, we derive the estimate of the LMS MSE for a complex sinusoidal interference and determine the optimum step-size parameter μ that results in minimum MSE. Then, we do the same for an AR(1) interference process, except in this case, the equalizer is implemented with the NLMS algorithm. Both interference scenarios can be interpreted as mathematical abstractions of a physical environment in which a high-data-rate communication signal is corrupted by a signal of much narrower bandwidth. We also include some numerical examples.

A. Sinusoidal Interference

In this scenario, the interference vector is given as shown in (32), shown at the bottom of the page, where ω_Δ is the offset frequency of the interference, and ϕ is a random phase uniformly distributed between $-\pi$ and π . The correlation matrix of the reference vector is given by

$$\mathbf{R} = (\sigma_s^2 + \sigma_n^2)\mathbf{I}_L + \sigma_x^2 \mathbf{v}\mathbf{v}^H \quad (33)$$

where \mathbf{I}_L is the $L \times L$ identity matrix. Then, using (6) and (7) and invoking the matrix inversion lemma [6, p. 565], the Wiener filter and corresponding MSE are given by

$$\mathbf{w}_w = \frac{1}{\sigma_s^2 + \sigma_n^2} \left(\mathbf{P} - \frac{\sigma_s^2 \sigma_x^2}{\sigma_s^2 + \sigma_n^2 + L\sigma_x^2} \mathbf{v} \right) \quad (34)$$

and

$$J_w = \frac{\sigma_s^2}{\sigma_s^2 + \sigma_n^2} \left(\sigma_n^2 + \frac{\sigma_s^2 \sigma_x^2}{\sigma_s^2 + \sigma_n^2 + L\sigma_x^2} \right). \quad (35)$$

The autocorrelation function of $u(k)$ is given by

$$r_u(m) = \sigma_s^2 \delta_m + \sigma_x^2 e^{i\omega_\Delta m} + \sigma_n^2 \delta_m \quad (36)$$

where δ_m is the Kronecker delta. Applying (36) to (25), we get

$$H_E(z) = \frac{e^{-i\omega_\Delta z} - 1}{e^{-i\omega_\Delta z} - 1 + \mu L \sigma_x^2}. \quad (37)$$

We cannot directly apply (27) and (28) because the discrete power spectrum $S_x(z)$ of a sinusoidal process is not defined. However, we can work in the frequency domain by applying the change of variable $z = \exp(i\omega)$ to (26)–(28) and using the fact that the power spectral density of the interference process is $S_x(e^{i\omega}) = 2\pi\sigma_x^2\delta(\omega - \omega_\Delta)$, where $\delta(\omega)$ is the Dirac delta. Then, we get

$$\begin{aligned} & |H_E(e^{i\omega})|^2 \\ &= \frac{2(1 - \cos(\omega - \omega_\Delta))}{1 + (\mu L \sigma_x^2 - 1)^2 + 2(\mu L \sigma_x^2 - 1) \cos(\omega - \omega_\Delta)} \end{aligned} \quad (38)$$

and

$$J_{\text{lms}} = \frac{1}{2\pi} \int_{-\pi}^{\pi} |H_E(e^{i\omega})|^2 \cdot (\sigma_s^2 |1 - W_w(e^{i\omega})|^2 + \sigma_n^2 |W_w(e^{i\omega})|^2) d\omega. \quad (39)$$

The sinusoidal interference does not contribute directly to the MSE because $H_E(e^{i\omega})$ has a zero at $\omega = \omega_\Delta$. The frequency response of the Wiener filter is

$$\begin{aligned} & W_w(e^{i\omega}) \\ &= \frac{\sigma_s^2}{\sigma_s^2 + \sigma_n^2} \left(1 - \frac{\sigma_x^2}{\sigma_s^2 + \sigma_n^2 + L\sigma_x^2} \sum_{l=-N}^N e^{-il(\omega - \omega_\Delta)} \right). \end{aligned} \quad (40)$$

Applying (38) and (40) to (39), the integral can be solved exactly. The steady-state LMS MSE is then given by

$$J_{\text{lms}} = \frac{2}{1 - A^2} (B - CA - DA^L) \quad (41)$$

where

$$\begin{aligned} A &= 1 - \mu L \sigma_x^2 \\ B &= C + D \\ C &= \frac{\sigma_s^2 \sigma_n^2}{\sigma_s^2 + \sigma_n^2} \\ D &= \frac{1}{\sigma_s^2 + \sigma_n^2} \left(\frac{\sigma_s^2 \sigma_x^2}{\sigma_s^2 + \sigma_n^2 + L\sigma_x^2} \right)^2 \end{aligned} \quad (42)$$

and $-1 < A < 1$. This condition on the variable A is required to make the filter $H_E(z)$ stable, although it may not be adequate to ensure stability of the LMS algorithm.

Next, we use (41) to find the step-size parameter μ_{opt} that results in the least MSE. Because A is the only term in (41) that depends on μ , we take its derivative with respect to A and determine that the A that minimizes this function is the zero of the $(L+1)$ -order polynomial

$$p(A) = C - 2BA + CA^2 + DLA^{L-1} + D(2-L)A^{L+1} \quad (43)$$

which is real and satisfies $-1 < A_{opt} < 1$. We make the additional approximation that the equalizer length L is large enough such that we can neglect the two highest order components of this polynomial and approximate it with the remaining quadratic equation. Then, the optimum step-size parameter that satisfies the stability condition can be written explicitly as

$$\mu_{opt} = \frac{1}{L\sigma_x^2} \left(1 - \frac{1}{C} (B - \sqrt{B^2 - C^2}) \right). \quad (44)$$

B. AR(1) Interference

In this scenario, the interference is generated by sending a complex, circular white Gaussian process through a filter with transfer function

$$H_{\text{AR}}(z) = \frac{z}{z - \alpha e^{i\omega_\Delta}}, \quad 0 < \alpha < 1. \quad (45)$$

The power of the Gaussian driving process is $\sigma_x^2(1 - \alpha^2)$. The autocorrelation function of the complex output AR(1) process is

$$r_x(m) = \sigma_x^2 \alpha^{|m|} e^{i\omega_\Delta m} \quad (46)$$

with discrete power spectrum

$$S_{\text{AR}}(z) = \frac{\sigma_x^2(1 - \alpha^2)z e^{-i\omega_\Delta}}{(ze^{-i\omega_\Delta} - \alpha)(1 - \alpha z e^{-i\omega_\Delta})}. \quad (47)$$

The adaptive equalizer is implemented with the NLMS algorithm because it is more likely to be stable in a region of the step-size parameter $\hat{\mu}$, where the performance improvement over the Wiener filter is most pronounced because it tends to average out short-term power fluctuations of the finite bandwidth AR(1) process and avoid possible gradient noise amplification [6, pp. 432–433].

$$\begin{aligned} \mathbf{x}^T(k) &= \sigma_x e^{i(\omega_\Delta k + \phi)} \begin{pmatrix} e^{i\omega_\Delta N} & e^{i\omega_\Delta(N-1)} & \dots & 1 & \dots & e^{-i\omega_\Delta(N-1)} & e^{-i\omega_\Delta N} \end{pmatrix} \\ &= \sigma_x e^{i(\omega_\Delta k + \phi)} \mathbf{v}^T \end{aligned} \quad (32)$$

The correlation matrix of the reference vector is written as

$$\mathbf{R} = (\sigma_s^2 + \sigma_n^2)\mathbf{I}_L + \mathbf{R}_{\text{AR}} \quad (48)$$

where \mathbf{R}_{AR} is the correlation matrix of the AR(1) interference process. An approximation of the Wiener filter is derived in the Appendix and is given by

$$\mathbf{w}_w(m) \approx \frac{\sigma_s^2 e^{i\omega_\Delta m}}{\sigma_s^2 + \sigma_n^2} \left(\delta_m + \frac{2ba - a^2 - 1}{1 - a^2} \cdot \left(a^{|m|} + (a^{-m} + a^m) \frac{a^L}{1 - a^L} \right) \right) \quad -N \leq m \leq N \quad (49)$$

where

$$\begin{aligned} a &= c - \sqrt{c^2 - 1} \\ b &= \frac{1 + \alpha^2}{2\alpha} \\ c &= b + \frac{\sigma_x^2(1 - \alpha^2)}{2\alpha(\sigma_s^2 + \sigma_n^2)} \end{aligned} \quad (50)$$

and $0 < a < 1$. Using (6) and (7), the MSE of the Wiener filter is approximated by

$$J_w \approx \sigma_s^2 \left(1 - \frac{\sigma_s^2}{\sigma_s^2 + \sigma_n^2} \left(1 + \frac{2ba - a^2 - 1}{1 - a^2} \cdot \left(1 + \frac{2a^L}{1 - a^L} \right) \right) \right). \quad (51)$$

Next, using (46) in (25), we get

$$R(z) = \sigma_x^2 \frac{\alpha e^{i\omega_\Delta}}{z - \alpha e^{i\omega_\Delta}} \quad (52)$$

and from (31)

$$\hat{H}_E(z) = \frac{ze^{-i\omega_\Delta} - \alpha}{ze^{-i\omega_\Delta} - \alpha \hat{A}} \quad (53)$$

where

$$\hat{A} = 1 - \frac{\sigma_x^2 \hat{\mu}}{\sigma_s^2 + \sigma_x^2 + \sigma_n^2}. \quad (54)$$

To ensure stability of the system $\hat{H}_E(z)$, we require $-1/\alpha < \hat{A} < 1/\alpha$.

We finally need the transfer function of the Wiener filter. To simplify the presentation, we assume that N is large enough such that all terms in the transfer function multiplied by N th powers of a or higher are neglected. Then, we get

$$W_w(z) \approx \frac{\sigma_s^2}{\sigma_s^2 + \sigma_n^2} \left(\frac{z^2 e^{-2i\omega_\Delta} - 2bze^{-i\omega_\Delta} + 1}{(ze^{-i\omega_\Delta} - 1/a)(ze^{-i\omega_\Delta} - a)} \right). \quad (55)$$

We then have the necessary components to apply (26) and (27) to (28). Although it is straightforward to solve (28) using the method of residues, it is cumbersome, and the general solution is quite complex. We therefore present solutions for three specific environments.

With a change of variables, it can be shown that we can set $\omega_\Delta = 0$ without loss of generality. The poles of the integrand of (28) that are inside the unit circle are at $z = 0$, $z = \alpha \hat{A}$, and $z = a$. Then, for $\sigma_s^2 = 2.0$, SNR equal to 25 dB (SNR = σ_s^2/σ_n^2), signal-to-interference ratio (SIR) equal to -20 dB (SIR = σ_s^2/σ_x^2), and $\alpha = 0.9999$, the steady-state NLMS MSE is given in (56), shown at the bottom of the page. As in the sinusoidal case, we find an \hat{A} that minimizes this function. The only optimum \hat{A} that meets the stability condition is $\hat{A} = 0.207761$. Using (54), we get $\hat{\mu}_{\text{opt}} = 0.800186$. In addition, because the transfer function of (53) and the approximate Wiener transfer function of (55) are independent of the equalizer length L , the theoretical expressions of the NLMS MSE and the associated optimum NLMS step-size parameter also are independent of the equalizer length.

For SNR = 15 dB and SIR = -20 dB, the steady-state NLMS MSE is given as in (57), shown at the bottom of the page, with $\hat{\mu}_{\text{opt}} = 0.428369$.

In addition, for SNR = 10 dB and SIR = -10 dB, the steady-state NLMS MSE is given as in (58), shown at the bottom of the page, with $\hat{\mu}_{\text{opt}} = 0.094487$.

C. Numerical Examples

To demonstrate the validity of these results, we present some numerical examples. We begin with the sinusoidal interference case. The theoretical MSE of the Wiener filter calculated using (35) and the analytical approximation of the LMS MSE determined from (41) are compared with the estimated LMS MSE derived experimentally by averaging over 50 realizations of the random processes. The communication signal $s(k)$ is

$$J_{\text{nlms}} = \frac{0.049555 - 0.103528\hat{A} + 0.063479\hat{A}^2 - 0.009505\hat{A}^3}{1 - 1.736660\hat{A} - 0.245805\hat{A}^2 + 1.736310\hat{A}^3 - 0.753844\hat{A}^4} \quad (56)$$

$$J_{\text{nlms}} = \frac{0.157575 - 0.401320\hat{A} + 0.336540\hat{A}^2 - 0.092794\hat{A}^3}{1 - 1.740060\hat{A} - 0.242849\hat{A}^2 + 1.739710\hat{A}^3 - 0.756800\hat{A}^4} \quad (57)$$

$$J_{\text{nlms}} = \frac{0.366835 - 1.066671\hat{A} + 1.033620\hat{A}^2 - 0.333779\hat{A}^3}{1 - 1.916320\hat{A} - 0.081725\hat{A}^2 + 1.915940\hat{A}^3 - 0.917891\hat{A}^4} \quad (58)$$

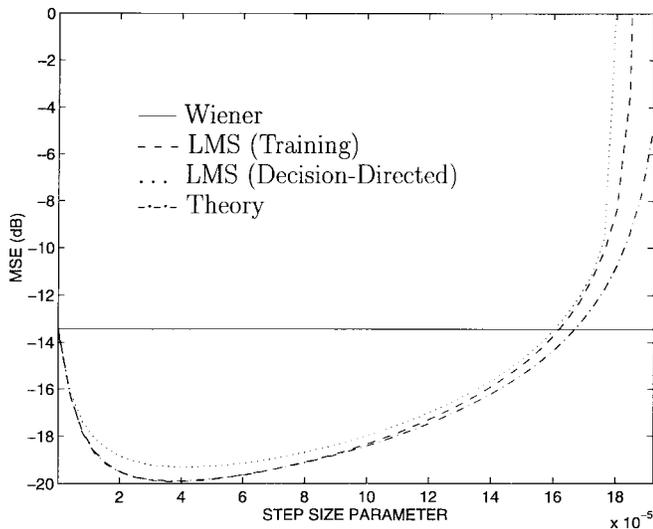


Fig. 2. LMS MSE as a function of μ for a sinusoidal interference with $L = 51$, SNR = 25 dB, and SIR = -20 dB.

simulated as a quadrature phase shift keyed (QPSK) signal in which the mutually independent in-phase and quadrature components take values $+1$ and -1 with equal probability ($\sigma_s^2 = 2$).

Fig. 2 is a plot of MSE as a function of the step-size parameter μ for equalizer tap length $L = 51$, SNR = 25 dB, and SIR = -20 dB. The estimated steady-state LMS MSE obtained during the training phase and the MSE obtained during the decision-directed mode are plotted. Clearly, the MSE performance improvement of the LMS algorithm over the Wiener filter can be significant with the proper choice of μ . Additionally, close agreement with theory is observed. The optimum choice of μ determined from (44) is $\mu_{opt} = 3.786676 \times 10^{-5}$, which also agrees with the figure and matches to eight significant digits the optimum step-size parameter calculated by including the two highest order terms in (43). Fig. 2 also is interesting because it contradicts conventional wisdom in adaptive filter theory in which a smaller step-size parameter μ is associated with less MSE. This is not the case for $\mu < \mu_{opt}$. However, as expected, the LMS MSE performance approaches that of the Wiener filter as $\mu \rightarrow 0$.

Fig. 3 is a plot of MSE as a function of the equalizer length L for SNR = 25 dB and SIR = -20 dB. The optimum LMS step-size parameter μ_{opt} is used. In addition, the MSE of the Wiener filter with a decision feedback filter is included. It is determined numerically using (15). The LMS algorithm exhibits less MSE than the corresponding Wiener filter and quickly approaches the performance of the Wiener filter with the decision feedback component as L increases, as surmised in Section II. In addition, because the approximation of (21) becomes better as L increases, close agreement between experiment and theory is observed for large L .

MSE as a function of SNR is shown in Fig. 4. The number of taps is $L = 51$, and SIR = -20 dB. Again, the step-size parameter is μ_{opt} . This figure demonstrates that for high SNR, the Wiener filter with decision feedback is able to cancel the interference without causing much distortion in the

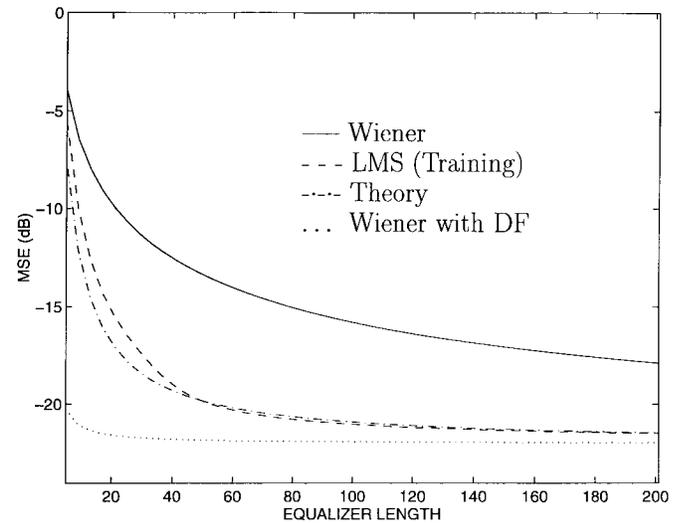


Fig. 3. LMS MSE as a function of equalizer length L for a sinusoidal interference using optimum step-size parameter μ_{opt} , SNR = 25 dB, and SIR = -20 dB.

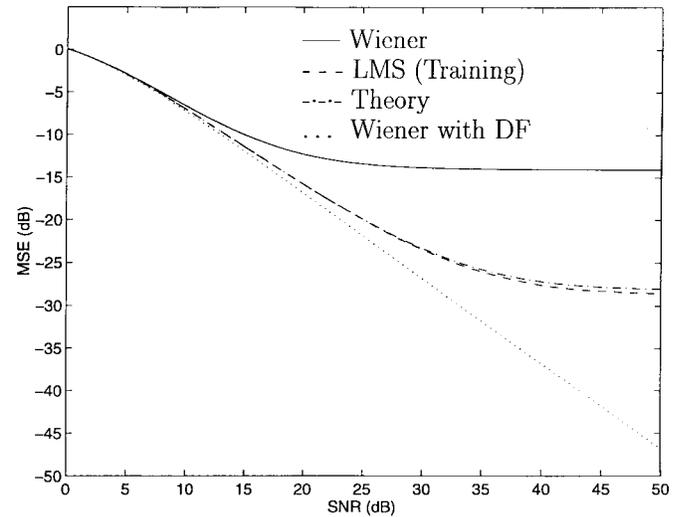


Fig. 4. LMS MSE as a function of SNR for a sinusoidal interference using optimum step-size parameter μ_{opt} , $L = 51$, and SIR = -20 dB.

communication signal and will outperform the feedforward transversal equalizer implemented with the LMS algorithm.

We next examine the AR(1) interference case when the equalizer is implemented with the NLMS algorithm. Fig. 5 is a plot of MSE as a function of $\hat{\mu}$ for the same environment used to generate Fig. 2. The Wiener MSE is calculated using (51), and the theoretical NLMS MSE results are determined using (56). We clearly see the degradation due to the finite bandwidth effects of the AR(1) process over the sinusoidal interference results of Fig. 2. Due to the emphasis placed on new data by a large value of $\hat{\mu}$, the equalizer in decision-directed mode becomes unreliable for large $\hat{\mu}$. The experimental MSE deviates from theory because the assumption made in the analysis that $\hat{s}(k) = s(k)$ is violated in this region of $\hat{\mu}$. In addition, the optimum $\hat{\mu}$ is seen to be approximately 0.8, as determined from (56).

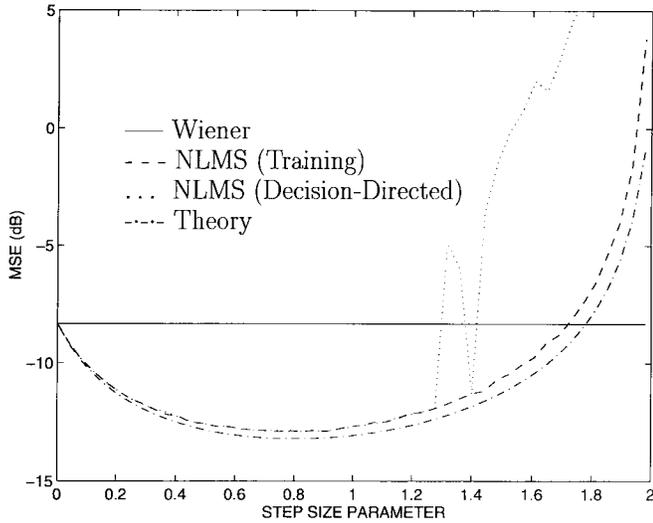


Fig. 5. NLMS MSE as a function of $\hat{\mu}$ for an AR(1) interference with $L = 51$, SNR = 25 dB, and SIR = -20 dB.

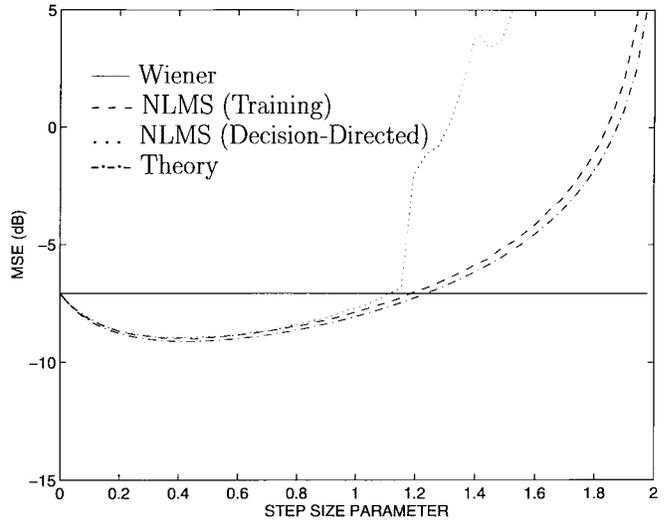


Fig. 7. NLMS MSE as a function of $\hat{\mu}$ for an AR(1) interference with $L = 51$, SNR = 15 dB, and SIR = -20 dB.

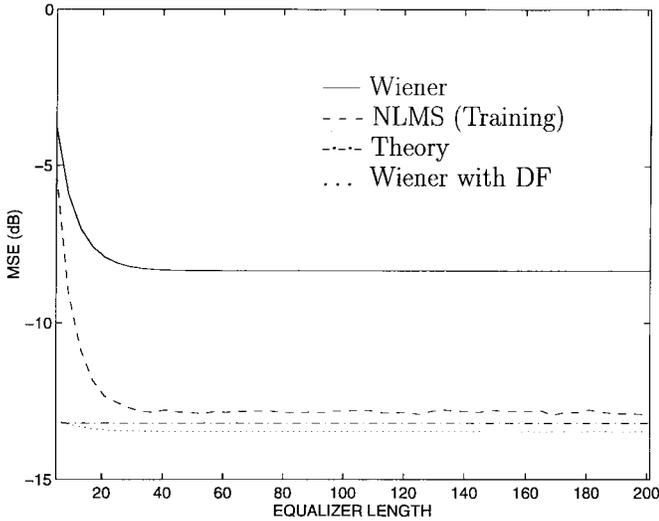


Fig. 6. NLMS MSE as a function of equalizer length L for an AR(1) interference using optimum step-size parameter $\hat{\mu}_{opt}$, SNR = 25 dB, and SIR = -20 dB.

Fig. 6 is a plot of MSE as a function of L for the environment used to generate Fig. 3. The NLMS algorithm is implemented with the optimum step-size parameter. It is evident that the theoretical approximation is good for $L > 20$. In addition, as determined from theory, MSE performance is relatively independent of L for the AR(1) interference.

Fig. 7 is a plot of MSE as a function of $\hat{\mu}$ generated using (57) for the scenario identical to that of Fig. 5, except with SNR = 15 dB. The MSE performance improvement of NLMS over the Wiener filter for this lower SNR case is not as dramatic as that depicted in Fig. 5. Finally, Fig. 8 represents the case for SNR = 10 dB with a much higher SIR of -10 dB using (58). Here, NLMS is seen to behave more in line with the conventional interpretation of adaptive filter performance. After an initial slight drop, the NLMS MSE increases as a function of $\hat{\mu}$. There also tends to be less agreement with theory as $\hat{\mu}$ increases. However, this region

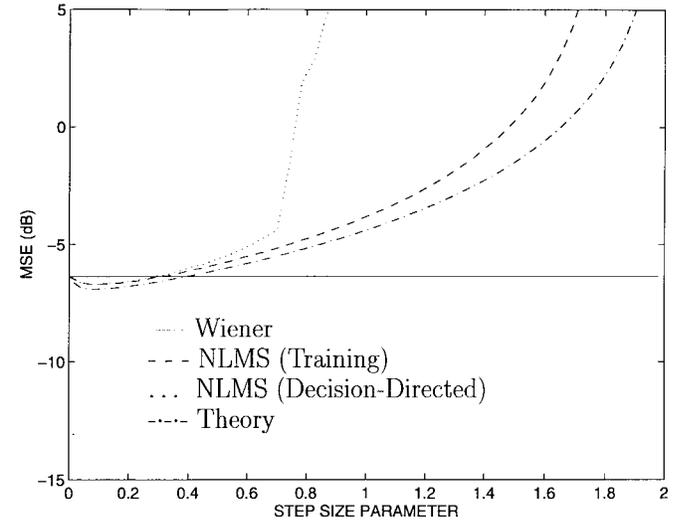


Fig. 8. NLMS MSE as a function of $\hat{\mu}$ for an AR(1) interference with $L = 51$, SNR = 10 dB, and SIR = -10 dB.

of $\hat{\mu}$ is of little practical importance because feedback error causes the equalizer in decision-directed mode to diverge.

V. CONCLUSIONS

The primary purpose of this paper is to demonstrate that an adaptive equalizer implemented with the LMS algorithm can have better narrowband interference canceling capabilities and exhibit better steady-state MSE performance than the corresponding Wiener filter. The LMS algorithm achieves this improvement in performance by effectively incorporating information not explicitly used by the Wiener filter. This phenomenon is approximated analytically without invoking the independence assumption by using a transfer function approach. This method uses only second-order moments of the processes and is independent of the underlying probability distributions. Although we have confined the presentation to an LMS-implemented adaptive equalizer structure, similar effects have been observed recently in other applications of

$$\frac{1}{\sqrt{L}} \left(e^{-i2\pi lN/L} \quad e^{-i2\pi l(N-1)/L} \quad \dots \quad 1 \quad \dots \quad e^{i2\pi l(N-1)/L} \quad e^{i2\pi lN/L} \right), \quad -N \leq l \leq N \quad (60)$$

adaptive filtering in which analytical measures derived using the independence assumption do not adequately quantify actual LMS performance such as in adaptive noise canceling [20] and adaptive system identification with temporally correlated additive noise [21].

APPENDIX

WIENER FILTER WITH AR(1) INTERFERENCE

To derive the Wiener filter for the AR(1) interference scenario, we assume that the number of taps L is large enough such that the Toeplitz matrix of (48) can be diagonalized approximately as

$$\mathbf{FRF}^H \approx \mathbf{\Gamma} \quad (59)$$

where \mathbf{F} is the unitary matrix with l th-row given as in (60), shown at the top of the page [22, pp. 141–145], and $\mathbf{\Gamma}$ is the $L \times L$ -diagonal matrix with elements given by

$$\begin{aligned} \mathbf{\Gamma}_{l+N+1, l+N+1} \\ &= \gamma_l \\ &= \sigma_s^2 + \sigma_n^2 + S_{\text{AR}}(e^{i2\pi l/L}), \quad -N \leq l \leq N. \end{aligned} \quad (61)$$

Then, using (6) and (7), we write the approximate Wiener filter weights as

$$\mathbf{w}_w \approx \frac{\sigma_s^2}{\sqrt{L}} \mathbf{F}^H \mathbf{\Gamma}^{-1} \mathbf{1} \quad (62)$$

where $\mathbf{1}$ is an $L \times 1$ -vector of ones.

Using (60), the approximate Wiener weights can be written as

$$\mathbf{w}_w(m) \approx \frac{\sigma_s^2}{L} \sum_{l=-N}^N \gamma_l^{-1} e^{-i2\pi ml/L}, \quad -N \leq m \leq N. \quad (63)$$

Assuming the interferer offset frequency ω_Δ is an integer multiple of $2\pi/L$ and using (47), we get

$$\mathbf{w}_w(m) \approx \frac{\sigma_s^2 e^{i\omega_\Delta m}}{L(\sigma_s^2 + \sigma_n^2)} \sum_{l=-N}^N \frac{b - \cos(2\pi l/L)}{c - \cos(2\pi l/L)} e^{-i2\pi ml/L} \quad -N \leq m \leq N \quad (64)$$

where

$$\begin{aligned} b &= \frac{1 + \alpha^2}{2\alpha} \\ c &= b + \frac{\sigma_x^2(1 - \alpha^2)}{2\alpha(\sigma_s^2 + \sigma_n^2)}. \end{aligned} \quad (65)$$

The summation in (64) can be solved by using the fact that

$$\begin{aligned} I(l) &= \int_{-\pi}^{\pi} \frac{b - \cos x}{c - \cos x} e^{-ilx} dx \\ &= 2b \int_0^{\pi} \frac{\cos lx}{c - \cos x} dx - 2 \int_0^{\pi} \frac{\cos lx \cos x}{c - \cos x} dx. \end{aligned} \quad (66)$$

Using tabulated solutions of these two integrals [23, pp. 366–367], we get

$$I(l) = 2\pi\delta_l + \frac{2\pi}{1 - a^2} (2ba - a^2 - 1)a^{|l|} \quad (67)$$

where $a = c - \sqrt{c^2 - 1}$. Using (65), and because $0 < \alpha < 1$, it can be shown that $0 < a < 1$. Then, using Poisson's sum formula [24, pp. 47–49], we get

$$\begin{aligned} \mathbf{w}_w(m) &\approx \frac{\sigma_s^2 e^{i\omega_\Delta m}}{\sigma_s^2 + \sigma_n^2} \frac{1}{2\pi} \sum_{l=-\infty}^{\infty} I(m + lL) \\ &\quad -N \leq m \leq N \\ &= \frac{\sigma_s^2 e^{i\omega_\Delta m}}{\sigma_s^2 + \sigma_n^2} \left(\delta_m + \frac{2ba - a^2 - 1}{1 - a^2} \right. \\ &\quad \left. \cdot \left(a^{|m|} + (a^{-m} + a^m) \frac{a^L}{1 - a^L} \right) \right). \end{aligned} \quad (68)$$

REFERENCES

- [1] S. U. H. Qureshi, "Adaptive equalization," *Proc. IEEE*, vol. 73, pp. 1349–1387, Sept. 1985.
- [2] J. D. Laster and J. H. Reed, "Interference rejection in digital wireless communications," *IEEE Signal Processing Mag.*, vol. 14, pp. 37–62, May 1997.
- [3] L. Li and L. B. Milstein, "Rejection of CW interference in QPSK systems using decision-feedback filters," *IEEE Trans. Commun.*, vol. COMM-31, pp. 473–483, Apr. 1983.
- [4] B. Widrow, J. M. McCool, M. G. Larimore, and C. R. Johnson, Jr., "Stationary and nonstationary learning characteristics of the LMS adaptive filter," *Proc. IEEE*, vol. 64, pp. 1151–1162, Aug. 1976.
- [5] W. A. Gardner, "Learning characteristics of stochastic-gradient-descent algorithms: A general study, analysis, and critique," *Signal Process.*, vol. 6, no. 2, pp. 113–133, Apr. 1984.
- [6] S. Haykin, *Adaptive Filter Theory*, 3rd ed. Englewood Cliffs, NJ: Prentice-Hall, 1996.
- [7] 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, pp. 31–39, Feb. 1979.
- [8] J. R. Treichler, "Transient and convergent behavior of the adaptive line enhancer," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-27, pp. 53–62, Feb. 1979.
- [9] N. J. Bershad and O. M. Macchi, "Adaptive recovery of a chirped sinusoid in noise—Part 2: Performance of the LMS algorithm," *IEEE Trans. Signal Processing*, vol. 39, pp. 595–602, Mar. 1991.
- [10] M. J. Shensa, "Non-Wiener solutions of the adaptive noise canceller with a noisy reference," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-28, pp. 468–473, Aug. 1980.
- [11] J. E. Mazo, "On the independence theory of equalizer convergence," *Bell Syst. Tech. J.*, vol. 58, no. 5, pp. 963–993, May/June 1979.
- [12] R. C. North, R. A. Axford, and J. R. Zeidler, "The performance of adaptive equalization for digital communications systems corrupted by interference," in *Proc. Asilomar Conf. Signals, Syst. Comput.*, Monterey, CA, 1993, pp. 1548–1553.
- [13] S. C. Douglas and W. Pan, "Exact expectation analysis of the LMS adaptive filter," *IEEE Trans. Signal Processing*, vol. 43, pp. 2863–2871, Dec. 1995.
- [14] H. J. Butterweck, "A steady-state analysis of the LMS adaptive algorithm without use of the independence assumption," in *Proc. IEEE ICASSP*, Detroit, MI, 1995, pp. 1404–1407.
- [15] J. R. Glover, "Adaptive noise canceling applied to sinusoidal interferences," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-25, pp. 484–491, Dec. 1977.
- [16] P. M. Clarkson and P. R. White, "Simplified analysis of the LMS adaptive filter using a transfer function approximation," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-35, pp. 987–993, July 1987.

- [17] J. G. Proakis, *Digital Communications*, 2nd ed. New York: McGraw-Hill, 1989.
- [18] D. L. Duttweiler, J. E. Mazo, and D. G. Messerschmitt, "An upper bound on the error probability in decision-feedback equalization," *IEEE Trans. Inform. Theory*, vol. IT-20, pp. 490–497, July 1974.
- [19] A. Papoulis, *Probability, Random Variables, and Stochastic Processes*, 2nd ed. New York: McGraw-Hill, 1984.
- [20] K. Quirk, J. R. Zeidler, and L. B. Milstein, "Bounding the performance of the LMS estimator for cases where performance exceeds that of the finite Wiener filter," in *Proc. IEEE ICASSP*, Seattle, WA, 1998, pp. 1417–1420.
- [21] H. J. Butterweck, "The independence assumption: A dispensable tool in adaptive filter theory," *Signal Process.*, vol. 57, no. 3, pp. 305–310, Mar. 1997.
- [22] A. K. Jain, *Fundamentals of Digital Image Processing*. Englewood Cliffs, NJ: Prentice-Hall, 1989.
- [23] I. S. Gradshteyn and I. M. Ryzhik, *Table of Integrals, Series, and Products*, 4th ed. San Diego, CA: Academic, 1980.
- [24] A. Papoulis, *The Fourier Integral and Its Applications*. New York: McGraw-Hill, 1962.



Michael Reuter (S'82–M'86) received the B.S. and M.S. degrees in electrical engineering from the University of Illinois, Urbana-Champaign, in 1984 and 1986, respectively. He is currently a Ph.D. candidate in the Department of Electrical and Computer Engineering, University of California, San Diego, La Jolla.

Since 1987, he has been an Engineer at the Space and Naval Warfare Systems Center, San Diego. His current research interests are in adaptive and statistical signal processing applied to problems in

wireless communications.



James R. Zeidler (M'76–SM'84–F'94) received the Ph.D. degree in physics from the University of Nebraska, Lincoln, in 1972.

Since 1974, he has been a Scientist at the Space and Naval Warfare Systems Center, San Diego, CA. He has also been an Adjunct Professor of electrical and computer engineering at the University of California, San Diego, La Jolla, since 1988. During this period, he has conducted research on communications systems, sonar and communications signal processing, communications signals exploitation, undersea surveillance, array processing, underwater acoustic communications, infrared image processing, and electronic devices. He was also a Technical Advisor in the Office of the Assistant Secretary of the Navy (Research, Engineering and Systems), Washington, DC, from 1983 to 1984.

Dr. Zeidler was an Associate Editor of the *IEEE TRANSACTIONS ON SIGNAL PROCESSING* from 1991 to 1994 and is a member of the Editorial Board of the *Journal of the Franklin Institute*. He received an award for the best unclassified paper at the IEEE Military Communications Conference in 1995 and the Navy Meritorious Civilian Service Award in 1991.