Performance Analysis of LMS Adaptive Prediction Filters

JAMES R. ZEIDLER, SENIOR MEMBER, IEEE

Real-time implementations of adaptive linear prediction filters have been shown to provide useful signal processing gains in a wide range of practical applications. The performance of such filters is generally bounded by that of the ideal Weiner filter, but the magnitude of the implementation errors is dependent on fixed adaptive filter parameters such as the adaptive time constant, the filter order, and the prediction distance. The goal of this paper is to delineate the conditions required to implement real-time adaptive prediction filters that provide nearly optimal performance in realistic input conditions. The effects of signal bandwidth, input signal-to-noise ratio (SNR), noise correlation, and noise nonstationarity are explicitly considered. Analytical modeling, Monte Carlo simulations and experimental results obtained using a hardware implementation are utilized to provide performance bounds for specified input conditions. It is shown that there is a nonlinear degradation in the signal processing gain as a function of the input SNR that results from the statistical properties of the adaptive filter weights. The stochastic properties of the filter weights ensure that the performance of the adaptive filter is bounded by that of the optimal matched filter for known stationary input conditions. Proper selection of the fixed filter parameters provide performance results which closely approximate that of the ideal Weiner filter.

I. INTRODUCTION

Linear prediction is a classic signal processing technique that provides estimates of the value of an input process at a time where no measured data are available based on measurements of the same process during intervals not containing the time of interest. The techniques are generally useful in separating signals from noise based on the differences in the signal and noise coherence times. They have been successfully applied to a wide range of applications including radar, sonar, spectral estimation, seismic processing, speech analysis/synthesis, communications, control theory, noise cancellation, system identification, astronomy, image processing, bioengineering, and other areas such as economic forecasting. Detailed formulations of the linear prediction problem and the mathematical procedures used to solve the resulting equations for a stationary stochastic process of interest are discussed in [1]-[5]. Recent advances in adaptive filtering algorithms and high speed digital hardware have produced families of computationally efficient algorithms which can be implemented in real time for many applications. The goal of this paper is to delineate the conditions required to implement real-time adaptive prediction filters that provide nearly optimal performance in practical applications.

Wiener was the first to derive an explicit formulation of the optimum least-squares predictor for stationary continuous-time stochastic processes in [6]. The related problem of defining the optimal linear least-squares filter for estimating a process which is corrupted by noise is also discussed [6]. It is shown that the coefficients of the optimal least squares predictor are defined in terms of the second order statistics of the process of interest. Consequently, determination of the optimal filter coefficients requires explicit knowledge of the autocorrelation or covariance functions of the input data. For discrete-time sampled data processes, the coefficients of a Wiener filter of order $L$ are defined by a set of $L$ linear equations known as the normal equations as described by Makhoul [1].

Efficient recursive methods for solving the normal equations for the optimal Wiener filter coefficients were developed by Levinson [7] and Durbin [8]. For a filter of order $L$, Durbin's method requires $L^2 + O(L)$ operations and $2L$ storage locations [1]. The solution of the normal equations is generally only a minor part of the computational load, however, since for an ensemble of $N$ data points, the computation of the covariance coefficients requires $LN$ operations, where typically $N >> L$ [1]. Equivalent procedures are utilized to define the optimal Wiener filter coefficients for locally stationary input processes as discussed by Makhoul [1]. Extensions of Wiener filter theory for nonstationary processes were made by Kalman [9] and others [5]. An extensive summary of linear filtering theory is provided by Kailath [5].

The design of Wiener filters requires a priori information about the statistics of the data to be processed. When such a priori information is unavailable, the required statistical information can be determined from the input data using least squares estimation. In this case the ensemble averages of the data which are used to obtain the Wiener filter coefficients are replaced by time averages of the input data over a defined time interval. The estimated statistical parameters are then inserted into the normal equations to obtain the optimal filter coefficients based on an assumption of sta-
tionarity over the defined time interval [1]-[3]. This procedure becomes cumbersome if the data is nonstationary since the process of estimating the second order statistics must be repeated frequently in order to track the variations in the input process [1], [3]. Block least squares and recursive least squares techniques have been utilized to provide a more efficient implementation in such cases [3]. Adaptive prediction based on gradient estimation are also frequently utilized to predict a time series with unknown statistical parameters [2]-[4].

Adaptive filters are self-designing based on a selected recursive or block algorithm which allows the filter to “learn” the initial input statistics and to track them if they vary slowly in time [2]-[4], [10]-[18]. These filters start from a predetermined set of initial conditions which contain no assumptions about the statistical characteristics of the data and then update the filter coefficients based on the chosen adaptive algorithm and the sequence of sampled data values. For stationary inputs it has been shown that properly designed adaptive filters will converge to the optimum Wiener filter within a factor known as the misadjustment noise [2]-[4], [10]-[18]. The misadjustment noise is determined by the magnitude of the errors inherent in the data adaptive estimation process, and can be controlled by appropriate selections of the fixed adaptive filter parameters. The rate at which the filter converges to the Wiener solution is defined by the learning time of the adaptive process (i.e., the adaptive time constant). The magnitude of the adaptive time constant and the misadjustment noise are inversely related [16]. Comparative performance of some of the adaptive algorithms which have been employed are provided in [2]-[4], [11]-[17].

The simplest and the most extensively utilized adaptive algorithm is the Widrow–Hoff Least Mean Squares (LMS) stochastic gradient algorithm [3], [10]-[22]. This algorithm utilizes a gradient search technique to determine the filter coefficients which minimize the mean square prediction error [12]. The LMS algorithm requires only 2N operations per iteration for real data (N for complex data [23]) and no explicit determination of the correlation coefficients of the input data [2]. Consequently these algorithms can offer significant computational advantages over Wiener filters.

Further computational reductions can be realized by the use of frequency domain adaptive algorithms which allow the time-domain convolutions of the time-domain algorithm to be replaced with multiplications by Discrete Fourier Transform (DFT) which can be implemented with Fast Fourier Transforms (FFT’s) [24]-[28]. For a filter of order L and a sequence of L data points, the total number of computations can be reduced from L² in the time domain to L log₂ L in the frequency domain [27].

The convergence rate for the LMS algorithm is determined by the eigenvalues of the input data covariance matrix. There is a separate time constant for each distinct eigenvalue of the covariance matrix [12]. The primary limitation of the LMS algorithm is the relatively long convergence times for those modes of the input data sequence which have small eigenvalues [2], [3], [24]-[31]. This limits performance for applications where large eigenvalue disparities in the input data sequence are encountered. One practical approach to mitigate this effect is to implement the adaptive filter in the frequency domain with sufficient frequency resolution to separate the primary components of the input process into separate frequency bins. As proven by Grenander and Szegö [32], as L becomes large the Toeplitz distribution theorem provides an eigenvalue decomposition of the input process. In this case each frequency bin can be updated independently using a normalized step parameter to equalize the convergence times despite the disparate input power levels at each frequency [28], [33]. This approach may only provide an approximate solution when the eigenspectra of the input data is unknown, however, since residual coupling between various components of the input data in the transform domain may exist [25], [26], [28], [33].

A more elegant approach to improve the adaptive filter convergence rate when there are large eigenvalue disparities in the covariance matrix of the input data is to utilize a lattice implementation [2]. It has been shown [2] that the Lth order lattice filter contains the prediction error filters of all lower orders, and that the prediction errors from each stage are mutually orthogonal. This can result in significantly improved convergence properties for lattice realizations of the Wiener filter [2], [3], [15]-[17], [29]-[31]. Another advantage of the lattice structure is that the filters are recursive in order and may achieve any order L, where L ≤ L, in order to optimize performance. The significant advantages of the lattice filter structure for speech synthesis and analysis is discussed by Makhoul et al. [15], [34]-[37].

Adaptive lattice filters based on the LMS gradient search algorithm were developed by Griffiths [38], [39]. A computationally efficient recursive implementation of an exact least squares lattice (LSL) was developed by Morf and coworkers [40], [41]. Various refinements to the recursive LSL algorithm are discussed in [2]-[4], [15]-[17], [42]. Alternate computationally efficient formulations of recursive least squares (RLS) adaptive transversal filters which utilize least squares estimation rather than stochastic gradient techniques were developed by Cioffi and Kailath [43] and by Carayannis, et al. [44]. The relative computational costs of various LSL and RLS algorithms is discussed in [2], [3], [15] (p. 217). It is shown that a number of RLS and LSL algorithms require only O(L) computations per iteration. The algorithms are thus significantly faster than the classical Wiener filter implementations, but the computational complexity can still be significant compared to the LMS algorithm because the order parameter is between 5 and 24 for the various realizations [2], [15]. In addition to the increased operation count of the LSL and RLS filters, the type of operations performed includes divisions and, in some cases, square roots in addition to the multiplies and adds required for LMS transversal filter. The performance advantages for a specific application must therefore be considered in terms of the increased computational complexity. Performance advantages of the LSL and RLS implementations are often significant, however, as discussed in [2]-[4], [14], [29]-[31].

Gentleman and Kung [45] developed an alternate structure to provide real-time solutions to the linear least-squares estimation problem using systolic array sections. In this case, [45], the input data matrix is orthogonalized by a pipelined sequence of Givens rotations. A simplified structure was developed by McWhirter [46]. Various implementations of adaptive filters using systolic array architectures are discussed in [3].

Milstein [47] discusses the difficulty of implementing many of the optimal Wiener filter structures for narrow-
band interference rejection in spread spectrum communications applications and stressed the use of practical receiver structures which provide nearly optimal performance in typical applications. Masry [48] provides closed form analytical results for applying linear predictive filtering to narrowband interference rejection in pseudonoise (PN) spread spectrum systems and notes the applicability of adaptive techniques such as those discussed by Anderson, et al. [49]. As discussed by Milstein and Itlits [50], the implementation of adaptive prediction filters in spread spectrum communication systems requires bandwidths of several hundred KHz for HF systems and more than 100 MHz in microwave systems. Real time implementations at these sampling rates is difficult to achieve. Consequently, the relative simplicity of the LMS algorithm remains an attractive feature in applications when the LMS algorithm is able to track the variations in the input statistics. In this case the initial delay required to achieve convergence may be an acceptable penalty for the reduced hardware complexity.

This paper will provide performance baselines for the LMS implementation of the adaptive prediction filter for a variety of stationary input conditions. Variations in filter performance which arise from changes in the spectral characteristics of the input signal and noise will be illustrated using analytical modeling, computer simulation, and a hardware realization of the LMS-adaptive prediction filter. Numerous papers have discussed the performance of various frequency estimator structures based on linear prediction at high signal-to-noise ratios (SNR). In this paper the filter characteristics are defined as a function of the input SNR and the fixed adaptive filter parameters. Comparative receiver operating characteristic (ROC) curves are utilized to compare the performance of a detection structure based on an ideal matched filter with that of the equivalent detection structure with an LMS-adaptive prediction filter in place of the matched filter. The probability density functions (pdf) and the ROC curves are provided for inputs consisting of sinusoids in white noise and for signals of finite bandwidth in white noise. It is shown that for proper selection of the fixed adaptive filter parameters, the performance of the adaptive detection structure corresponds closely to that of the optimal matched filter previously derived by Urkowitz [51]. The effects of signal bandwidth and noise correlation on the characteristics of the Wiener and the LMS-adaptive prediction filter are discussed. Extensive Monte-Carlo simulation results on the pdfs of the adaptive filter weights and adaptive filter output are provided to illustrate the performance bounds imposed by real-time implementation.

Although the useful applications of adaptive filters are for nonstationary or for unknown input conditions, the analytical results obtained for these specified stationary cases illustrate limitations of the LMS-adaptive prediction filter relative to the optimal Wiener filter. The magnitude of the differences between the optimum and adaptive filter solutions are computed as a function of the selectable filter parameters (i.e., the filter order, the adaptive feedback constant, and the prediction distance). The results identify the parameter settings required to maximize performance in various applications. It is also shown that detection and estimation structures which are based on the output statistics of the adaptive LMS filter are often superior to those based on the LMS filter weights.

One of the most interesting results obtained is the non-linear degradation in output SNR as a function of input SNR. This results from the statistical properties of the adaptive filter weights. The adaptive prediction filter is shown to provide large signal processing gain at high SNR at the consequence of producing asymmetrical probability density functions with significant energy in the tails of the distribution. This ensures that the detection performance of both the Wiener and the LMS adaptive prediction filter are bounded by classical detection theory. In nonstationary noise backgrounds it is shown that the LMS-adaptive prediction filter significantly reduces the effect of the noise fluctuations when compared to the conventional matched filter. This can provide improved performance for nonstationary noise environments.

II. ADAPTIVE LINE ENHANCEMENT

An adaptive implementation of a linear prediction filter which was updated using the LMS algorithm was first proposed by Widrow et al. [10], [52] to provide a real-time capability for many situations. This configuration was designated as the Adaptive Line Enhancer (ALE), due to the initial application of separating sinusoids from broadband noise. It is generally applicable, however, whenever there is a significant difference between the correlation times of the signal and the noise. Consequently, the ALE and related adaptive prediction filter implementations have been investigated for a wide range of applications including instantaneous frequency estimation, spectral analysis, and narrowband detection [10], [53]-[54]; speech encoding [54], [65]; narrowband interference rejection [47]-[50], [66]-[74]; predictive deconvolution [77]; intrusion detection [78], [79] and adaptive carrier recovery for digital data communications receivers [80].

Several distinct implementations of ALE have also been used. Finite impulse response (FIR) time domain algorithms, using either the adaptive filter weights [10], [53], [55], [59], [78], the adaptive filter output [52], [62], [80]-[82] or the prediction error signal [69] were first employed. Subsequent implementations included infinite impulse response (IIR) transversal filters [79], [83]-[91], adaptive FIR lattice filters [61], [92]-[94], IIR lattice filters [80], [95], frequency domain FIR filters [27], [28], [96], block LMS implementations [25], [97], and multi-rate implementations [98].

Recent results with improved IIR LMS adaptive filter implementations [80], [99] indicate that the problems encountered with filter instability for the first IIR adaptive filter realizations can sometimes be satisfactorily controlled, but it remains difficult to insure robust performance in many applications [80], [84]-[91], [99]-[103]. As noted by Cupo and Gitlin, the ALE-VLA provides two major advantages over the IIR-ALE: stability during adaptation and a unimodal performance surface. Fan and Jenkins [99] have also observed that for echo cancelling applications the IIR adaptive filter can provide superior performance after convergence, but showed that the IIR filter required a longer convergence time and was more sensitive to proper selection of the filter order. An IIR-ALE structure which incorporates a lattice structure was shown by Cupo and Gitlin [80] to provide superior performance to an FIR-ALE for certain cases in removing phase jitter noise in a digital data communications receiver. The analysis in this paper is restricted to the FIR-ALE implementation.
The steady-state performance of the ALE implemented as a Δ-step linear predictor for multiple sinusoids in white noise was first analyzed in [55]. The performance of the ALE for detection of multiple sinusoids in uncorrelated noise and the estimation of the frequencies of the embedded sinusoids is addressed in [55], [56], [61], [93]. Yorgamanadanum, et al. [56] points out that although there are a number of other techniques which have been applied to this problem, the ALE is of interest because of its simple structure. It was shown by Zedler, et al. [55] that a variation of the prediction distance, Δ, allows improved frequency estimation performance for multiple sinusoids in white noise. It is also shown [49], [62], [63], [92] that improved performance is also achieved for Δ > 1 for narrowband signals in correlated noise when compared to a conventional auto-regressive (AR) time-series analysis techniques with Δ = 1. The analytical results derived in [55] were shown to be in good agreement with the results obtained with a LMS-hardware implementation.

An optimal prediction distance for the estimation of the sinusoidal frequencies at high input SNR was derived by Reddy et al. [61]. Gupta [68] showed that the optimal Δ for high SNR was valid for an arbitrary SNR in the case of a single sinusoid in white noise. Gupta [68] further proved that the mean-square error of the optimized ALE (neglecting misadjustment noise) is equal to that of the optimized two-sided Wiener filter for this case. Yorgamanandum [56] extended these results to obtain an optimal value of Δ for the case of two sinusoids in white noise for an arbitrary SNR. It was shown that the use of the ALE with optimally chosen Δ provides frequency estimates which have small bias and variance which is close to the Cramer–Rao bound above a critical threshold SNR [56]. Below this threshold SNR, these estimates were shown to degrade rapidly as a function of input SNR. In this paper, the factors which define this SNR threshold will be delineated.

Bershad and Qu [104] have shown that the optimum log likelihood ratio test statistic for a detector configuration based on the LMS adaptive filter weights is almost equivalent to that of an optimal detector configuration for low input signal-to-noise ratios (SNR), but is significantly degraded from the optimal detector for high input SNR. In this paper it will be shown that this degradation in performance at high SNR is a property of the Wiener filter solution and can often be minimized by basing the detector test statistic on the adaptive filter output rather than the adaptive filter weights.

In order to compare performance of various ALE implementations and to provide accurate performance bounds at low SNR, it is essential to include the weight misadjustment noise [18]. A steady-state performance model which allows the relative contributions of the Wiener filter and the misadjustment terms to be compared as a function of the input spectral characteristics is provided below. The time required to reach steady-state is also discussed.

III. ALE PERFORMANCE MODEL

A. Steady-State Performance

As illustrated in Fig. 1, the time domain FIR ALE consists of an L-weight linear prediction filter in which the coefficients w(k) are adaptively updated at the input sampling rate, f_s. The ALE output, r(k), is defined by

$$r(k) = \sum_{\ell=0}^{L-1} w_\ell(k) x(k - \ell - \Delta)$$  \hspace{1cm} (1)

where Δ is the prediction distance of the filter in units of the sample period. The output, r(k), is subtracted from the input sequence, x(k), to form a prediction error sequence, ε(k), where

$$\varepsilon(k) = x(k) - r(k).$$  \hspace{1cm} (2)

This error sequence is scaled by a multiplication of magnitude 2μ and fed back to adjust each of the L filter weights according to the Widrow–Hoff LMS algorithm:

$$w(k+1) = w(k) + 2\mu \varepsilon(k) x(k - \ell - \Delta),$$  \hspace{1cm} \hspace{1cm} \hspace{1cm} \ell = 0, \ldots, L - 1$$  \hspace{1cm} (3)

where $\mu$ is preselected based on the desired performance characteristics. The adaptive algorithm can be expressed in vector form in terms of the input data vector, x(k), i.e.,

$$w(k+1) = w(k) + 2\mu \varepsilon(k) x(k - \Delta),$$  \hspace{1cm} (4)

where,

$$x(k - \Delta) = [x(k - \Delta), x(k - 1 - \Delta), \ldots, x(k - L + 1 - \Delta)]^T.$$  \hspace{1cm} (5)

![Fig. 1. Block diagram of the adaptive line enhancer (ALE).](image-url)
A steady-state model for the converged ALE weight vector \( \mathbf{w}(k) \triangleq [w_0(k), \cdots, w_{L-1}(k)]^T \) was obtained by Rickard and Zeidler [105] by decomposing \( \mathbf{w}(k) \) into two parallel weight vectors, i.e., \( \mathbf{w}(k) = \mathbf{w}^* + \mathbf{w}(k) \), as shown in Fig. 2. Here \( \mathbf{w}^* \) is the optimum Wiener \( \Delta \)-step predictor weight vector defined by [1]

\[
\mathbf{w}^* = R_{xx}^{-1} \mathbf{d},
\]

(6)

where \( R_{xx} \) is the \( L \times L \) autocorrelation matrix of the input vector \( x(k) \) and \( \mathbf{d} \) is the \( L \times 1 \) cross-correlation vector between the primary input, \( x(k) \), and the reference channel input vector, \( \mathbf{w}(k-\Delta) \). \( \mathbf{w}(k) \) is a very slowly fluctuating, zero-mean, vector random process due to wide noise. Under the assumptions of relatively slow adaptation, and stationary input and noise, the statistics of weight vector "misadjustment" \( \mathbf{w}(k) \) are virtually constant over a reasonable observation time [71], [105]–[108]. Thus \( \mathbf{w}(k) \) may be treated as a vector random variable \( \mathbf{w} \), that is independent of time, with mean \( \mathbf{w}^* \) and covariance \( \mathbf{w}^T \), where

\[
\mathbf{w} = \mathbf{w}^* + \mathbf{w}(k) \quad \text{in (7)},
\]

with \( \mathbf{I} \) the identity matrix and \( \mathbf{w}^* \) is the minimum mean-squared error associated with Wiener filtering of the input [105],

\[
\mathbf{w}^* = \mathbf{R}_{xx}^T \mathbf{R}_{xx}^{-1} \mathbf{d}.
\]

(7)

This simplified performance model treats the ALE output (which is in general a nonlinear function of the input data and the adaptive filter parameters, \( \mu, L \), and \( \Delta \)) in terms of the mean values of the converged adaptive filters weights and the statistical properties of the random displacements of the converged weights from their mean values. This model has been shown to provide an accurate representation of the adaptive filter performance for a wide range of input conditions [46], [71], [82]. These results are extended by Bershad et al. [106]–[108] to larger values of \( \mu \) where the time varying properties of the weights give rise to intermodulation products of the input and the filter weight fluctuations. The time required to achieve steady state performance is defined below.

\[\text{B. Mean Convergence Rates of the LMS Adaptive Filter Weights and Output}\]

The convergence properties of LMS adaptive filters are robust under a variety of adverse conditions, provided only

that the inputs to the filter satisfy specific nondegeneracy conditions [11], [17], [109], [110]. For many applications the presence of background noise over the entire information band will ensure that the LMS algorithm will converge in the mean to a stable solution. The LMS algorithm is as efficient as any constant feedback algorithm can be, provided that all the eigenvalues of the data autocorrelation matrix are equal [18]. In general, however, the LMS adaptive learning curves consist of a sum of modes associated with each distinct eigenvalue and the misadjustment noise is determined primarily by the largest eigenvalues, while the convergence rates are defined by the smallest eigenvalues [12].

Treichler [57], [58] first analyzed the convergence properties of the ALE for a sinusoid in white Gaussian noise (WGN) using the assumption that \( R_{xx} \) is an autocorrelation matrix for a real-valued stationary stochastic input process. For this case \( R_{xx} \) must be positive, semidefinite, symmetrical, and Toeplitz with a set of orthonormal eigenvectors. Shensa [111], [112] utilized asymptotic properties of Toeplitz matrices such as the Toeplitz distribution theorem derived by Grenado and Szegő [29] to define the ALE time constants and learning curves for multiple sinusoids in WGN and for a narrowband signal in WGN.

Shensa [111], [112] has shown that the mean-squared difference between the expected value of the ALE weight vector at time \( k, \mathbf{w}(k) \), starting from an initial value of zero, and the optimal Wiener–Hopf solution, \( \mathbf{w}^* \), is given by

\[
\| \mathbf{w}^* - \mathbf{w}(k) \|^2 = \sum_{n=1}^{L-1} (1 - 2\mu \lambda_n)^2 \mathbf{E}^\ast \cdot \mathbf{w}^* |^2,
\]

(9)

(9)

In (9) \( \lambda_n \) and \( \mathbf{E}^\ast \) are the eigenvalues and eigenvectors of the input data autocorrelation matrix, \( R_{xx} \), and \( \mathbf{E}^\ast \cdot \mathbf{w}^* \) is the scalar projection of the \( n \)-th eigenvector onto the optimal Wiener filter weight vector. A similar expression is derived for the convergence of the mean-square prediction error for the case where the gradient estimation noise is neglected (this noise is of order \( L \mu \mathbf{K}^2 \)) [111]:

\[
\mathbf{E} - \mathbf{E}^\ast = \sum_{n=1}^{L-1} (1 - 2\mu \lambda_n)^2 \mathbf{E}^\ast \cdot \mathbf{w}^* |^2,
\]

(9)

where \( \mathbf{E} = \mathbf{E}(k) \) is the mean square prediction error power at time \( k \), and \( \mathbf{E}^\ast \) indicates the optimal steady-state mean square value. Note that the learning curves for both the
weights and the error are composed of $l$ terms that relax geometrically, with a constant logarithmic slope of log $(1 - 2\mu\lambda)^3$ which is approximately equal to $4\mu\lambda$. Thus, each term grows to within $e^{-1}$ of its optimal value in a time given by:

$$\tau_n = \frac{1}{4\mu\lambda}. \tag{11}$$

Consequently, the maximum modal time constant $\tau_{max} = (4\mu\lambda_{max})^{-1}$ is a conservative estimate of filter performance, since only those eigenvalues for which the projection of the final state onto the corresponding eigenvector (i.e., $E^* w^*$ in (9) or $\lambda_i |E^* w^*|^2$ in (10)), is large will exert significant influence in convergence time. Since the eigenvalue $\lambda_i$ may be small even in cases where $|E^* w^*|$ is not negligible, the adaptive filter error convergence may be controlled by even fewer modes than the adaptive filter weights (i.e., $\lambda_i |E^* w^*| \ll |E^* w^*|$ for the modes where $\lambda_i \ll 1$). Consequently, the filter output often converges more rapidly than the filter weights.

1) Convergence for Multiple Sinusoids in WGN: Shensa [111], [112] derived expressions for the convergence rates of the ALE weights for an input of $M$ sinusoids in WGN in the form

$$x(k) = n(k) + \sum_{n=0}^{M} A_n \sin(\omega_n kt + \theta_n) \tag{12}$$

where $n(k)$ is a WGN sequence with uniform noise power, $\nu^2$. Shensa [111] shows that the $L$ dimensional vector space of the input may be decomposed into an $M$-dimensional subspace spanned by the signal component and its orthogonal complement. Consequently, the optimal ALE weight vector lies within the signal subspace and all eigenvectors outside this space are orthogonal to $w^*$, and hence their eigenvalues do not affect the learning curves in (9) and (10). Consequently, only $2M$ non-zero scalar projections of $w^*$ onto the eigenvectors of the $R_\alpha$ matrix exist for this case, and for $0 \ll \omega_i \ll \pi$ the two eigenvectors corresponding to the component of frequency $\omega_i$ are equal. Consequently, there are at most $2M$ modes that define the convergence rates of an $L$-weight ALE for the input defined in (12). Further, for the above conditions and $M = 1$, the ALE learning curves exhibit a single time constant defined by:

$$\tau_n = [\mu(A_0^2 L + 4\nu^2)]^{-1}. \tag{13}$$

Also, for $M > 1$, if $(\omega_i - \omega_j) < 2\pi/L$ for all $n$ and $i$, the learning curves exhibit an independent mode for each distinct sinusoid. If the independent sinusoids have equal power, the eigenvalues are equal so that there is only a single mode in the learning curve. If the signals have unequal power, but are independent, then there will be a mode for each distinct signal power $A_n$ as defined in [13]. Treichler [57], [58] also analyzed the decay time for the filter to forget a signal which suddenly disappears after the filter has converged. In this case,

$$\tau_n = (4\mu\lambda^3)^{-1}, \text{ for all } n. \tag{14}$$

Consequently, the decay rate may be significantly longer than the learning time.

When $(\omega_n - \omega_j) < 2\pi/L$, interference between the individual sinusoids can produce an appreciable slowing of the convergence of the filter weights. Shensa [112] has analyzed this case in detail for a closely spaced frequency doublet and has shown that the convergence rates of the two non-zero convergence modes depend on the sum and difference frequencies of $\omega_j$ and $\omega_n$. For this case there will be two distinct modes, even when both sinusoids have equal power.

2) Convergence Rates for Narrowband Signals in WGN: Using the asymptotic properties of Toeplitz matrices for large $L$ derived in [32], Shensa [111], [112] also derived the convergence properties of the ALE for a finite bandwidth signal in WGN. The narrowband signal model considered was a one-pole, complex narrowband signal in WGN with an input autocorrelation function, $R_{xx}$, defined by

$$R_{xx}(t) = a_0^2 \exp\{-\alpha t + i\omega_0 t\} + a_0^2 \delta(t). \tag{15}$$

The eigenvectors of $R_{xx}$ are indexed by the frequency, and the $n$th eigenvalue corresponds to the power spectral density at $\omega_n$. Consequently, the modes of the learning curves are also indexed by frequency, and the fastest converging modes of the ALE weights occur at the frequencies $\omega$ for which the power spectral density of $x(k)$ is a maximum (i.e., for $\omega - \omega_0 \leq \alpha$ in (15)). Consequently, the evolving transfer function of the LMS adaptive filter will initially be narrow centered about $\omega_0$, and will widen with increasing time as the slower modes at frequencies away from $\omega_0$ begin to converge. This is illustrated by computer simulation results in Fig. 3 [111], [112]. The filter output was shown to converge much more rapidly than the filter weights because of the relatively smaller amount of signal energy present in these slower modes as defined in (9) and (10). Shensa [111], [112] verified these theoretical results using computer simulations to determine the relative convergence rates of the ALE weights and the ALE output.

3) Tracking Performance for Nonstationary Input Signals: ALE performance with transient frequency shifts between two tonal frequencies $f_2$ and $f_1$ was analyzed by Anderson and Shensa [113]. It was shown that for the special case when both $f_2$ and $f_1$ are integral multiples of $L$, the ALE response to the signal shift is an independent combination of two components: the decay from the previous state with a time constant defined by (14) and the convergence to a new state with time constant defined by (13) with $A_n$ defined by the signal power of $f_1$. In all other cases there is a coupling between the old and the new frequency components which accelerates the convergence of the weight vector to a new Wiener filter solution for the signal at frequency $f_1$. The magnitudes of the cross-coupling between the old and the new states of the solution are determined by the magnitudes of the scalar projections of the old and new eigenvectors onto the old and new Wiener filter solutions as defined in [113]. The normalized weight vector error at time $k$, $|N|^2 w^* - w(k)|$, was computed [113] as a function of the frequency shift, $\Delta f = n/2L$, and the time in filter lengths. The results obtained for four different input signal-to-noise ratios is shown in Fig. 4. The maximum weight vector error was shown to occur for $\Delta f = f_2 - f_1 = n/2L$, where $n$ is any integer. The convergence time increases rapidly for $\Delta f \gtrsim 0.15f/2L$, but also decreases rapidly with increasing SNR. The precipitous initial decrease in the weight vector error at $-10$ dB SNR with subsequent slow asymptotic decrease results from the mixture of a very rapid convergence toward the new state $w^*_1$ combined with a much slower decay from the old state $w^*_0$. 

1786

PROCEEDINGS OF THE IEEE, VOL. 78, NO. 12, DECEMBER 1990
ALE response to FM chirped and doppler shifted sinusoids was analyzed by Treichler [59] and Bershad et al. [60], [114]-[116]. Other treatments of LMS adaptive filter response to nonstationary signal conditions are provided in [117]-[121]. Convergence properties of adaptive lattice filters are developed in [122]-[125].

Bershad et al. [60] analyzed the transient behavior of the ALE for a complex sinusoid with linearly varying frequency in additive noise. It was shown that for a given input sweep rate, $\alpha$, the amplitude of the weight response depends on $\mu$, SNR, $\alpha$, and $L$, but that the weight scaling does not alter the relative phase of the weights. In order to minimize the mean square error power, the adaptive filter must match the signal phase at the output of the ALE to that of the input. As the frequency rate of change increases, the end weights of the filter must introduce large changes in the time-delayed signal phase. This is most easily done when the weight amplitudes are small. Consequently the weight amplitudes are attenuated with increasing weight index if

---

**Fig. 3.** Power spectrum of $W(k)$ versus time, for $k = 100, 250, 500,$ and $\infty$ iterations.

---

**Fig. 4.** Weight vector error for a transient frequency shift in tonal frequency. (To facilitate comparisons of the error surfaces for different SNRs, the vertical scale is normalized by dividing computed values of (6) by the factor $4\pi^2/(12)$. The normalized error is plotted as a function of time and as a function of frequency shift for $L = 2048$, $\mu = 1.2 \times 10^{-4}$, $\sigma^2 = 0.25$, $\Delta = 1$, $\omega_c/\omega_c = 3/2$ for the four designated values of SNR. The frequency shift is measured in units of $L^{-1}$. (a) SNR = $-40$ dB, (b) SNR = $-30$ dB, (c) SNR = $-20$ dB, (d) SNR = $-10$ dB.

---

ZEIDLER: LMS ADAPTIVE PREDICTION FILTERS

---

1707
the sweep rate exceeds certain conditions specified in [60]. It is shown [60] that the mean square error increases for increasing \( L \) for cases where the end weights of the filter have no significant response to the input. As the sweep rate increases for a fixed observation time, \( T \), a larger value of \( \mu \) is required to track the input in order to obtain minimum mean square error.

IV. OUTPUT POWER SPECTRAL DENSITY FOR SINUSOIDAL SIGNALS IN WGN

Consider an input to the ALE given by

\[
x(k) = A \sin (\omega_0 k + \theta) + n(k)
\]

where \( \omega_0 \) is known, \( \theta \) is arbitrary, and \( n(k) \) is a zero-mean white Gaussian sequence of variance \( \sigma^2 \). It was shown in [56] that the elements of the mean ALE weight vector \( \mathbf{w}^* \) are given by

\[
\mathbf{w}^* = \frac{2a^*}{L} \cos \left[ \omega_0 (l + \Delta) \right], \quad 0 \leq l \leq L - 1,
\]

where

\[
a^* = \frac{2SNR}{L} \sqrt{\frac{SNR}{2}},
\]

and

\[
SNR = A^2/2\sigma^2.
\]

It was further shown in [55] that the ALE steady-state response to \( N \) sinusoids in WGN can be expressed as a linear superposition of \( N \)-independent ALE’s, each adapted to a single sinusoid in WGN, provided that the separation of the center frequencies, \( \Delta \omega \), significantly exceeds the ALE frequency resolution (i.e., \( \Delta \omega > > 2\pi/L \)), and that \( 0 << \omega_n << 2\pi \) for all \( \omega_n \). The frequency resolution capabilities of both real and complex implementations of the ALE and the optimal choice of the bulk delay parameter \( \Delta \) to eliminate biases in the frequency estimates for cases of multiple sinusoids in WGN when \( \Delta \omega < \) are discussed in [55], [56], [93].

It was shown in [105] by Rickard and Zeidler that the adaptive filter output, \( r(k) \), is a scaled sinusoid embedded in non-Gaussian noise of zero mean, \( n_r(k) \), i.e.: \( r(k) = a^* A \sin (\omega_0 k + \theta) + n_r(k) \).

It was further shown [105] that, for the usual case when \( \Delta_{\text{min}} = \pi^2 \), the power spectral density of \( r(k) \), \( P_r(\omega) \), is expressed by:

\[
P_r(\omega) = \frac{\pi A^2}{2} \frac{a^* \sigma^2}{L} \left[ \sin^2 \left( \frac{\omega - \omega_0}{L} \right) + \sin^2 \left( \frac{\omega + \omega_0}{L} \right) \right] \mu \sigma^2 L \]

\[
+ a^* \sigma^2 \left[ \frac{1 - \cos \left( \frac{\omega - \omega_0}{L} \right)}{L^2} \left[ \frac{1 + \cos \left( \frac{\omega - \omega_0}{L} \right)}{1 + \cos \left( \frac{\omega - \omega_0}{L} \right)} \right] \right]
\]

\[
\text{for } -\pi < \omega < \pi.
\]

The four distinct components of \( P_r(\omega) \) result from convolving the respective signal and noise components of the input with the parallel Wiener filter and misadjustment filter model components defined above. These components are illustrated in Fig. 5. The two sinusoidal components of \( P_r(\omega) \) result from convolving the input sinusoid with the Wiener filter and the misadjustment filter. A broadband noise term is produced by convolving WGN with the misadjustment filter. The narrowband filtered noise component results from convolving the input WGN with the Wiener filter. The ALE output spectrum is thus composed of a sinusoid at the center of a noise pedestal of narrow band filtered noise, which is embedded in a broadband noise spectrum. The relative magnitudes of the noise pedestal and the broadband noise depend nonlinearly on \( L, \mu \), and the input SNR, as illustrated in [105]. It is also shown [105] that the mean value of the ALE output, \( \dot{r}(k) \), averaged over the ensemble of weight vectors is given by:

\[
\dot{r}(k) = a^* A \sin (\omega_0 k + \theta).
\]

Consequently the mean power of the output sinusoid will match the power of the input sinusoid only when \( L \cdot \text{SNR} \rightarrow 1 \) so that \( a^* \rightarrow 1 \). Further \( \dot{r}(k) \rightarrow 0 \) as \( A \rightarrow 0 \). By increasing the frequency resolution of the FIR adaptive filter, however, the relative magnitudes of the output sinusoidal power and the narrowband filtered noise power are increased relative to the misadjustment components. The output SNR of the ALE was defined [105] in terms of the input SNR as the ratio of the power spectrum density of the deterministic component of \( r(k) \) relative to the power spectral density resulting from the three random components in (20). Expressions for the narrowband gain in a specified bandwidth are also provided in [105]. In order to obtain experimental verification of these results, simulation results for the output power spectral density are described below.

From (20), the output power spectral density at \( \omega_0 \) is given...
by
\[ P_{r}(\omega) = \left( \frac{\pi A^2}{2} \delta(\omega - \omega_0) + \nu^2 \right) (a^2 + \mu^2 L). \] (22)
Likewise, for a frequency \( \omega \), where the narrowband filtered noise term is negligible, the output power spectral density reduces to
\[ P_{l}(\omega) = \mu^2 L, \quad \text{for } |\omega - \omega_0| \gg 2\pi/L. \] (23)
Since the input power spectral density is given by
\[ P_{l}(\omega) = \frac{\pi A^2}{2} \left[ \delta(\omega - \omega_0 + \delta(\omega + \omega_0)) + \nu^2 \right], \] (24)
\[ \frac{P_{r}(\omega)}{P_{l}(\omega)} = \mu^2 L \quad \text{for } |\omega - \omega_0| \gg 2\pi/L. \] (25)
Likewise,
\[ \frac{P_{r}(\omega)}{P_{l}(\omega)} = (a^2 + \mu^2 L) \] (26)
Thus, from (25), for \(|\omega - \omega_0| \gg 2\pi/L\), the effect of the ALE is to reduce the power spectral density of the input noise by the factor \( \mu^2 L \). Further since \( a^2 \leq 1 \) and typically \( \mu^2 L \ll 1 \), the output power spectral density at \( \omega_0 \) is also reduced, but approaches unity at values of \( L \) and SNR such that \( a^2 \rightarrow 1 \). This is illustrated in the simulation results of Figs. 6(a)-(c) for a 2048 weight ALE with \( \mu = 1.5 \times 10^{-3} \) and input SNRs of -25 dB, -30 dB, and -35 dB.

In Figs. 6(a)-(c), the relative magnitudes of the power spectral densities \( P_{r}(\omega) \) and \( P_{l}(\omega) \) are obtained by a spectral analysis of the ALE input and output. It is shown that for \(|\omega - \omega_0| \ll 2\pi/L\), the power spectral density of the input noise is reduced by \( 30 \pm 1 \text{ dB} \) for all input SNRs. This is in agreement with the noise reduction predicted from (25), i.e., \( \mu^2 L = -29.1 \text{ dB} \) for the case illustrated in Figs. 6(a)-(c). An expression for the ALE gain can be obtained by comparing the ratio of the input and output power spectral densities at \( \omega_0 \) to the ratio at \( \omega \) for \(|\omega - \omega_0| \ll 2\pi/L \), i.e.,
\[ C(\omega) = \frac{P_{r}(\omega)}{P_{l}(\omega)} \quad \text{for } |\omega - \omega_0| \gg 2\pi/L, \] (27)
From (25) and (26),
\[ C(\omega) = \left( 1 + \frac{a^2}{\mu^2 L} \right). \] (28)
The measured gain as a function of input SNR for \( \mu = 1.5 \times 10^{-3} \) and \( L = 2048 \) is compared to that calculated from (28) in Fig. 7. It is shown that good agreement between theory and experiment is obtained.

As indicated in Fig. 6, the ALE output noise spectrum exhibits a higher variance than the input noise spectrum. The statistical properties of the ALE weights and output are addressed below. The above results illustrate that the threshold SNR at which performance degrades is a function of both \( \mu \) and \( L \). As discussed in [126] however, an optimum value of \( \mu \) is also a function of the amount of available data. An optimal value of \( \mu \) is obtained in section VI below for a fixed observation interval.

V. ALE PERFORMANCE FOR NARROWBAND SIGNALS

A. Parameter Estimation for Narrowbands Signals in WGN

Analytical expressions for the impulse response and transfer function of a Wiener linear prediction filter and for

Fig. 6. Input and output power spectral density: \( L = 2048 \), \( \mu = 1.5 \times 10^{-3} \). (a) Input SNR = -25 dB, (b) Input SNR = -30 dB, (c) Input SNR = -35 dB.

Fig. 7. Comparison of measured and calculated gain in output power spectral density vs. input SNR: \( L = 2048 \), \( \mu = 1.5 \times 10^{-3} \).
the output power spectral density of the ALE were derived by Anderson et al. [49] for inputs composed of one or more stationary narrowband signals in WGN with a stationary autocorrelation function defined by (15). The signals were composed of WGN passed through a filter of center frequency \( \omega_0 \) with bandwidth \( \alpha \) that was small relative to the Nyquist frequency. It was shown [49] that the optimal Wiener filter takes the form of a bandpass filter, with bandwidth \( \beta \), where \( \beta \) was shown to exceed the bandwidth of the input signal \( \alpha \), except for very low input SNR. \( \beta \) was shown [49] to be a function of the input SNR and input signal bandwidth, according to the relationship:

\[
\cosh \beta = \cosh \alpha + \frac{\text{SNR}}{2} \sinh \alpha, \quad \alpha, \beta << 1, \quad L >> 1
\]

(29)

The Wiener filter bandwidth, \( \beta \), is plotted as a function of input SNR for various values of \( \alpha \) in Fig. 8. The vertical bar on each curve indicates the point where SNR = \( \alpha \).

![Fig. 8. Wiener filter bandwidth \( \beta \) versus input SNR for several values of signal bandwidth \( \alpha \). The small vertical bar on each curve indicates the point where SNR = \( \alpha \).](image)

Subsequently, estimates of signal bandwidth based on the ALE weights exhibit a positive bias at high SNR.

The relative magnitudes of the Wiener filter components and the misadjustment noise terms which arise from the gradient noise of the LMS algorithm are also derived by Anderson et al. [49] using the ALE steady state performance model defined in Section III-A above. The relative magnitude of the four primary components of the ALE output power spectral density for \( \alpha = 1/L \), \( L = 1024, \mu = 2^{-13} \), and \( \text{SNR} = -30 \text{ dB} \) is shown in Fig. 10. For this case, the Wiener

![Fig. 9. Frequency response of Wiener filter for several values of SNR where \( \mu = 2^{-13}, \Delta = 1, \alpha = 2^{-13} \equiv -15 \text{ dB}, L = 16/\alpha = 512 \). Small vertical bars mark \( \omega = \omega_0 \pm \beta \), the 3-dB filter bandwidth.](image)

![Fig. 10. An example of power spectral densities of the four ALE output components for the parameters \( \mu = 2^{-13}, \Delta = 1, L = 1024, \alpha = 1/L, \text{SNR} = -30 \text{ dB} \). (a) WF signal and noise terms. (b) MF signal and noise terms.](image)
filter signal component and the Wiener filter noise component are of equal amplitude at $\omega = \omega_0$ because the ratio of the power spectral density of the signal to that of the noise is approximately unity. The misadjustment filter (MF) terms are more than 20 db lower than the Wiener filter (WF) terms at $\omega = \omega_0$ but the white misadjustment noise component dominates all other components for $(\omega - \omega_0) \geq 0.025\pi$.

The optimal ALE length and bulk delay parameters for maximum SNR enhancement were derived in [49] as a function of the input bandwidth and SNR. For the narrowband signal model considered, the broadband ALE gain was shown to decrease approximately exponentially as a function of increasing values of the product $\alpha\Delta$. This effect results from decorrelation of the finite bandwidth signal. It was shown that the ALE bulk delay should not exceed $1/\alpha$ when the input noise is uncorrelated. The choice of optimal filter length was shown to depend upon input SNR as well as signal bandwidths as illustrated in Fig. 11. Under low SNR conditions, where the misadjustment terms of (20) dominate the ALE output noise, $L_{opt} \approx 1/\alpha$. For higher SNR, where the Wiener filter noise component is significant, maximum gain is typically achieved for a filter length shorter than the inverse signal bandwidth. The major conclusion is that attempting to over-resolve narrowband signals by increasing $L$ eventually degrades ALE gain because of larger misadjustment noise at the output [49].

B. Parameter Estimation for Narrowband Signals in Correlated Noise

The ALE response for sinusoidal signals in correlated noise was analyzed in [62], [63], [92], [127]. In this section some of the key results obtained for correlated noise will be verified experimentally. The mean steady-state impulse response and the transfer function of the LMS adaptive filter will be examined for three types of stationary inputs: one-pole low-pass noise, a narrowband signal embedded in one-pole low-pass noise, and a sinusoid in bandpass noise.

1) Low Pass Noise: For one-pole low-pass filtered input noise with an autocorrelation function given by $R_v(k) = P_0 \exp(-\alpha k)$, the Wiener filter component of the ALE impulse response is shown by Satorius et al. [62], [127] to be nonzero only at the first weight, with a magnitude that decreases exponentially with increasing delay according to the expression

$$w^*(k) = e^{-\alpha k}b(k). \quad (30)$$

This is verified by hardware simulation in Fig. 12 for three different values of $\Delta$, for $\alpha = \pi/7$ and $L = 256$. The first weight is shown to decrease as the ratio $e^{-\alpha} : e^{-2\alpha} : e^{-3\alpha} = 2.45 : 1.57 : 1$, which is in agreement with (30).

2) Narrowband Signal in Low Pass Noise: The second case considered is an input consisting of a narrowband signal centered at frequency $\omega_1$ embedded in additive one-pole low pass noise. The signal was generated by filtering uncorrelated noise with a two-pole bandpass filter. The bandwidth of the signal was small enough that the sinusoidal approximation discussed in [49] is valid. The expression for the Wiener filter impulse response was derived by Satorius et al. [127] and shown to be

$$w^*(k) = A_1 e^{i\omega_1 k} + A_2 e^{-i\omega_1 k} + C_1^* \delta(k - r + 1) + C_1^* \delta(k + r - 1). \quad (31)$$

Fig. 11. Optimal filter length determination as a function of signal bandwidth and SNR. (a) $\omega = 2^{-15}$, $\Delta = 1$, small vertical bars indicate $L = 1/\alpha$ for each curve. Note that the length yielding maximum gain is equal to $1/\alpha$ for $SNR \geq \alpha$, but $L_{opt} < 1/\alpha$ for $SNR \approx \alpha$. (a) $SNR = -50$ db, (b) $SNR = -30$, (c) $SNR = -10$ db.

The constants $A_1$, $A_2$, and $C_1^*$ were derived in [117] for the case where coupling between $A_1$ and $A_2$ can be ignored. It was shown that:

$$A_1 = \bar{A}_2 \approx \frac{e^{i\omega_1} - e^{-\alpha d}}{L + 2/SNR + (\cos \omega_1 - e^{-\alpha})(\cosh \alpha - \cos \omega_1)} \quad (32a)$$

$$C_1^* = e^{-\alpha d} - \sum_{n=1}^{\alpha} \frac{\bar{A}_n}{1 - e^{-\alpha + j\omega_1}} \quad (32b)$$
Another characteristic of the Wiener filter which can be observed in Fig. 13 is the increase in the amplitude of the sinusoidal component of the weight vector response for $\Delta = 4$ due to the increased decorrelation of the input noise. The ratio of the measured values of the amplitude of the sinusoidal component of the impulse response in Fig. 13 is given by 0.5:1. The predicted value for this ratio (calculated from (32) with $\Delta = 1$ and $\Delta = 4$) is also given by 0.5:1.

The noise decorrelation property of the $\Delta$-step adaptive prediction filter is apparent in this case since the large impulse at the beginning of the filter gives rise to an all-pass component in the filter’s frequency response. For small values of $\Delta$, the Wiener filter will pass a considerable amount of noise due to this all-pass component, thereby degrading the response to the signal components.

**Sinusoid in Bandpass Noise:** Another example which illustrates the noise suppression capabilities that can be achieved by varying the prediction distance of the Wiener filter is presented in Fig. 14. In this case, the steady state frequency response of the ALE is plotted for four different values of the delay when the input consists of a sinusoid embedded in additive bandpass noise. The noise was generated by filtering white noise through an eight-pole Butterworth low-pass filter followed by an eight-pole Butterworth high pass filter. For purposes of comparison, the
input spectrum is also plotted in Fig. 14. As indicated, the bandpass noise is suppressed as $\Delta$ is increased but the sinusoid is not attenuated for any value of $\Delta$ selected. Further note that significant decorrelation of the input noise occurs as $\Delta$ is increased to $\Delta = 20$, but little additional decorrelation occurs between $\Delta = 20$ and $\Delta = 100$, since the 60 Hz bandpass noise is essentially decorrelated after 1/60 sec.

The Wiener filter will treat broadband inputs as "signals" for values of $\Delta$ that are small compared to the reciprocal bandwidth. This is shown in Fig. 15 using an ALE hardware implementation. The input and output of the ALE are shown for a delay of 0.29 ms in Fig. 15 for the input noise spectrum used in Fig. 14 above. The power spectral density of the error signal and the ALE weights are shown in Fig. 16 for this same condition. The spectral density of the ALE output provides an estimate of the ALE transfer function, and the power spectral density of the error signal indicates the spectral regions in which coherent energy was removed. Note that the ALE input and output powers were approximately equal and that the 17 db reduction achieved in the noise power in the error signal was equal to the quantization noise limit of the log quantized ALE hardware used in these tests.

The above results illustrate that the optimal value of $\Delta$ for noise suppression lies in an interval between the correlation distance of the noise and the correlation distance of the signal. A much tighter bound on $\Delta$ is necessary to
achieve optimal resolution of multiple frequency components in the input if the frequencies are closely spaced [56], [57], [61], [93].

VI. COMPARATIVE DETECTION PERFORMANCE OF ADAPTIVE AND CONVENTIONAL DETECTORS

Several distinct types of adaptive detector structures have been proposed for use in environments where insufficient a priori information on the signal and noise statistics is available to design fixed optimal detectors. Detector structures that utilize adaptive mean level detection algorithms, sequential detection, nonparametric detection, pattern recognition, and decision-directed feedback techniques are discussed in numerous references [128]-[140] for a variety of signal and noise environments. Adaptive detectors are typically ad hoc designs that are based on knowledge of some general features of the signal and noise, but are capable of adjusting their internal structure through observation of the received data. Because their structure is data-dependent, a detection performance analysis is generally much more difficult than for nonadaptive detectors.

The conventional detector for the detection of narrow-band signals in WGN is implemented as a bank of narrow-bandpass complex (in-phase and quadrature) filters whose passbands span the frequency range of interest, as shown in Fig. 17. The complex output of each filter is sampled at the end of a fixed observation interval and quadratically detected to form the basic detection statistic for each passband. The statistic is optimal in the Neyman-Pearson sense for detecting a sinusoid of known frequency centered in a given passband, and of unknown phase in WGN of known power [141]. The above system may be implemented using the fast DFT algorithm to perform the complex filtering operation [51], [142]. The detection statistics for this case are the set of magnitude-squared DFT coefficients corresponding to the frequency range of interest. If the narrowband signal duration exceeds the coherent integration time of the DFT, then the detected filter outputs for a given passband may be incoherently integrated to form an improved statistic [52], [143]. The composite filter overlap requirements needed to attain optimal performance for unknown signal center frequency is discussed in [142]. Optimal detection in colored Gaussian noise of unknown covariance is addressed in [137]-[140].

The ALE has been proposed for use as a prefilter to the conventional DFT detector described above [10], [64], [78], [81], [82], [139], [143]. The use of the ALE as a prefilter to the conventional matched filter detector was shown by Nielson and Thomas [139] to provide improved detection performance in nongaussian nonstationary Arctic ocean noise. Rickard, et al. [82], [143] also illustrated the performance improvements which the ALE prefilter can provide for the detection of a narrowband signal in nonstationary noise. Two distinct ALE detection structures have been proposed: the ALE output detector and the ALE weight detector.

Statistical properties of the ALE weights and output are defined in [71], [81], [82], [104]. The ALE weight detector implementation was first proposed in [10] and analyzed by Kriyes in [64]. As discussed in [64], detection is based on a single L-point DFT of the ALE weight vector after the ALE has processed N data points, where \( L \ll n \). The detection statistic is formed from the L-point DFT. Postdetection integration is not employed in this implementation. For the ALE output detector, a magnitude squared \( K \)-point DFT is performed on the output, \( r(k) \), and then integrated incoherently over a time interval, \( T_r \).

The complexity of the ALE transfer functions for bandpass signals in white noise derived in [49] and the complexity of the output noise statistics derived in [81] for sinusoid in WGN indicates that a detailed analysis of the ROC performance of ALE-augmented detectors for finite bandwidth signals is a formidable task. Rickard, et al. [82] provided analytical results for the ALE/DFT detector with no incoherent integration for sinusoids in WGN. In this section, the available analytical results and extensive experimental results obtained using Monte-Carlo simulation for both conventional matched filter detectors and ALE-augmented detectors for inputs consisting of a finite bandwidth signal in WGN will be discussed. The simulation results are shown to substantiate the analytical results and to provide parametric descriptions of ALE/DFT performance for cases in which analytical results are not obtainable. The goal of this work is not to define the structure of the optimal detector for nonstationary inputs, but rather to delineate how the preselected ALE parameters affect detection performance. ROC curves are obtained for a range of signal bandwidths, filter lengths, prediction distances, adaptive time constants, and postdetection integration times, \( T_r \), of the ALE-augmented processor. These ROC curves are evaluated to determine: a) the statistical properties of the ALE output; b) the best choice of the ALE parameters for various input conditions; c) the sensitivity in performance to that choice; and d) the comparative performance of the two distinct adaptive detector implementations that have been proposed. The performance of each ALE-augmented detector implementation is compared to that of conventional detectors for known input conditions.

![Fig. 17. DFT detection processor with postdetection integration.](image-url)
A. Simulation Configuration

The difficulty encountered in the ROC analysis of the ALE/DFT detector arises from the complicated joint statistical properties of the ALE output, \( r(k) \), which is a convolution of a stochastic process with a set of random filter weights. This necessitated the development of an extensive computer simulation facility to verify the analytical models and to examine other detection properties which are not amenable to detailed analysis. The Rockwell International Sonar Simulation facility developed by M. Dentino was utilized for this purpose. The Monte Carlo simulations were completed at Rockwell International [143]–[146] by passing either white noise or a signal and white noise through both the ALE-augmented and conventional detection systems as illustrated in Fig. 18. Simulation results compiled included the probability density functions (pdf) of the ALE-augmented detector output as a function of the postdetection integration time constant, \( T_r \), the moments of the pdf's, and comparative ROC curves for various input conditions. For this study, the total processing time interval \( T \) and \( f_s \) were held constant and \( \alpha \), SNR, \( l \), \( \tau_a \), \( \Delta \), and \( T_r \) were varied. In all cases \( T = 128 \) seconds, \( f_s = 1024 \) Hz, the input noise bandwidth was 512 Hz, and the signal center frequency was bincentered with respect to the output FFT. The FFT resolution, \( \beta_i \), and the resolution of the FIR ALE, \( \beta_1 \), is given by:

\[
\beta_i = f_s / K; \quad \beta_1 = f_s / L.
\]

The input consisted of a sampled zero-mean white Gaussian process summed with a narrowband signal. The narrowband signal was generated by passing a zero-mean WGN process through a single-pole filter and translating the filtered output to the desired frequency. The response time of the filter, the frequency of the carrier, the signal bandwidth and the power of the resulting narrowband process were all program controllable. The outputs of the two ALE-augmented systems were sampled at every system response time, \( T_r \) and processed by the appropriate statistical analysis algorithms to determine the pdf's, the moments of the pdf's, and the ROC curves. A simulation of the conventional FFT processor was also performed simultaneously as shown in Fig. 18. An exponentially weighted postdetection integrator was used for both the ALE output detector and conventional detection processors. The integrated outputs for both processors were obtained by rectangular windowing of the data without overlapping the time series.

For the noise-only condition, at least 500 independent test samples were used to generate noise-only pdf's and to determine the threshold required for a given probability of false alarm, \( P_{fa} \). A total of at least 630 independent samples was made at each input SNR for the signal-to-noise input condition. The signal-plus-noise simulations were repeated for a range of SNRs. The associated probability of detection, \( P_d \), as a function of the threshold setting which is required to achieve the desired \( P_{fa} \), was obtained for each SNR. A relatively small number of samples for the signal-plus-noise input conditions was adequate for the determination of intermediate values of \( P_d \).

B. Statistical Properties of the Adaptive Filter Weights and Output

The probability of an adaptive filter weight exceeds a given threshold \( \psi \) for a WGN input was measured by using Monte Carlo simulation. The resultant \( P_{fa} \) is plotted in Fig. 19(a) as a function of the threshold (measured in normalized units of the variance of the pdf) for \( L = 1024 \) and \( \tau_a = 85 \) s. The simulation results for the adaptive filter weight, \( \phi_{i}(\psi) \), are shown to be in excellent agreement with a Gaussian pdf.

The discrete Fourier transform of the \( L \) filter weights at the frequency \( f_s \) is defined by a sum of \( L \) complex, zero-mean Gaussian variables and is also a Gaussian random variable. Also plotted in Fig. 19(a) is the measured probability function, \( \phi_{i}(\psi) \), for the real component transform of the weights, \( \text{Real} \{ W_i \} \) (the imaginary component is equivalent to the real component). Again, there is excellent agreement between the measured and theoretical results. The magnitude squared transform of the filter weights, \( \phi_{i}(\psi) \) is shown to have a two degree of freedom chi-square density function as shown in Fig. 19(a).

The filter output, \( r(k) \) is defined by a convolution of the

---

**Fig. 18.** Simulation configuration of ROC determination.
filter weights with the input data vector. Since \( w_I(k) \) is an independently distributed zero-mean Gaussian random variable (0, \( \sigma_w^2 \)) for an input zero-mean Gaussian variable (0, \( \sigma^2 \)), the density function of a single term in the summation of (1) is given by

\[
P(z) = \frac{1}{\pi \sigma_w^2} K_0 \left( \frac{|z|}{\sigma_w} \right)
\]

(35)

where \( z = w_I(k) x (k - \ell - \Delta) \) and \( K_0(\cdot) \) is a zeroth-order Bessel function of the third kind. For large \( L \), the density function for the output of an \( L \)-weight filter approaches a Gaussian function. The experimental probability function \( \phi_3(\cdot) \) for \( L = 1024 \) is seen in Fig. 19(b) to be consistent with a Gaussian density. The underlying non-Gaussian behavior of the filter output becomes apparent, however, when the transform of the output is taken as shown in [82].

**C. ROC Performance for Sinusoids in WGN of Known Power**

The ALE/DFT detector without postdetection integration is evaluated below under the two test hypotheses \( H_0 \) and \( H_1 \), where \( H_0 \) implies that the signal is absent and \( H_1 \) implies that the signal is present.

1) \( H_0 \) Case (Signal Absent): Under \( H_0 \), the input to the system \( x(k) \) is a white Gaussian noise sequence with variance \( \sigma_n^2 \), and thus \( w^* = 0 \). The elements of the ALE weight vector are independent, identically distributed (i.i.d.), zero-mean Gaussian r.v.'s with variance \( \sigma_w^2 = \mu \sigma_n^2 \), as described above. Given a particular realization \( w \) of the weight vector, the ALE output \( r(k) \) is a conditional Gaussian process. We denote this conditioning by \( r(k|w) \). Now let

\[
R_{jiw} = \sum_{k=0}^{K-1} r(k|w) \exp \left(-j \frac{2\pi k}{K} \right)
\]

(36)

be the \( j \)th DFT coefficient, where \( 0 \ll j \ll (K/2) - 1 \), and define \( y_{jiw} = |R_{jiw}|^2 \), which is the basic test statistic. Note from (7) that, conditioned on \( w \), \( R_{jiw} \) is a circular Gaussian r.v., and thus \( y_{jiw} \) has a gamma distribution with two degrees of freedom (dof) [1]. Dropping the subscript notation, the pdf of \( y \) is given by

\[
p(y|w) = \frac{1}{\sigma_y^2} \exp \left(-\frac{y}{\sigma_y^2} \right)
\]

(37)

where \( \sigma_y^2 = \mathbb{E}_w \{ y \} \) is the conditional expectation of \( y \) with respect to the input process \( x(k) \), i.e., \( \sigma_y^2 \) is still a function of \( w \). Since \( p(y|w) \) depends on \( w \) only through the scalar r.v. \( \sigma_y^2 \), we may write \( p(y|w) = p(y|\sigma_y^2) \).

Let \( p_0^2(u) \) denote the pdf of \( \sigma_y^2 \), and \( p_0(y) \) denote the desired unconditional pdf of \( y \). By the rule of total probability, \( p_0(y) \) is given by

\[
p_0(y) = \int_0^\infty du \ p(y|u)p_0^2(u)
\]

\[
= \frac{2\sigma_y^{n+1}}{\Gamma(n+1)} y^{n/2} K_n(2\sqrt{uy}).
\]

(38)

where \( K_n(\cdot) \) denotes the modified Bessel function of the third kind.

Given a threshold \( \alpha \), the false-alarm probability \( P_{FA} \) for \( y \) may be computed from (38) as

\[
P_{FA} = \int_0^\infty dy \ p_0(y) = \frac{2(\alpha a)^{n+1/2}}{\Gamma(n+1)} K_{n+1}(2\sqrt{\alpha a}).
\]

(39)

Thus (38) and (39) completely specify the ALE/DFT detector performance for the \( H_0 \) case, with no incoherent integration.

2) \( H_1 \) Case: With a sinusoidal signal present at the input, the mean ALE weight vector \( w^* \) is nonzero. Thus the ALE output contains four distinct, uncorrelated components, corresponding to signal and noise filtered by the mean (Wiener) weight vector and misadjustment weight vector, respectively, as illustrated in Fig. 5. This multiplicity of out-
put components renders intractable the method of analysis used for the $H_0$ case. However, the relative magnitudes of the Wiener and misadjustment filter transfer functions in the neighborhood of the sinusoid frequency permits the simple but accurate approximation which was utilized in [B2] to obtain results that correspond with the theoretical results.

D. Determination of Optimal Adaptive Filter Parameters

1) Optimal Time Constants: The total response time of an ALE output detector is the sum of the adaptive time constant of ALE and the response time of the post detection integrator. To ensure a valid comparison of the performance of the conventional and ALE output detectors, the total response time of the adaptive detector was matched to the response time of the postdetection integrator in the conventional detector, i.e.,

$$T = r_A + T_S. \quad (40)$$

Decreasing $r_A$ thus has the effect of decreasing $P_{FA}$ due to the increased smoothing of the noise-only probability distributions by the postdetection integration process over the interval $T_S$. If $r_A$ is made too small, however, the increased misadjustment noise will dominate the smoothing effects of the postdetection integrator. This trade-off in performance is illustrated in Fig. 20(a) and 20(b) for various choices of $r_A$.

Extensive simulations were performed to determine the optimal ratio of $r_A/T$ for $T = 128 \text{ s}$ for the ALE output detector. Fig. 21 indicates that for a sinusoidal input signal, the optimal time constants occur at $r_T/T = 2/3$ for a range of input SNRs. Equivalent results were obtained for various $P_{FA}$ in Fig. 22 and for sinusoidal input signals. Plots of $P_D$ versus $r_A/T$ for a range of input signal bandwidths as shown in Fig. 23. Again, the same optimal value of $r_A = (2/3)T$ is obtained. Consequently for the conditions simulated

$$r_{opt} = (2/3)T. \quad (41)$$

2) Sensitivity to Prediction Distance: As discussed above, the ALE SNR enhancement is achieved when the signal correlation time exceeds the noise correlation time. For signals embedded in WGN, any selection of $\Delta \geq 1$ will decorrelate

$$\text{Fig. 20.} \quad \text{False-alarm probability as a function of adaptive time constant for exponential postdetection integration with } (r_A + T_S) \text{ fixed at 128 sec.}$$

$$\text{Fig. 21.} \quad \text{Sensitivity to time constant for various SNRs.}$$

$$\text{Fig. 22.} \quad \text{Sensitivity to time constant for various } P_{FA}.$$
the background noise. As the noise correlation increases, progressively larger values of $\Delta$ are required. In Fig. 24 the $P_D$ for a fixed SNR of $-30$ dB and a fixed $P_{FA}$ of $10^{-4}$ is plotted as a function of $\Delta$ for various signal bandwidths. As indicated, the detection loss due to signal decorrelation increased for increasing $\Delta$ as the signal bandwidth is increased for $\alpha > f_s/2L$. For $\alpha \leq f_s/2L$, no appreciable loss in $P_D$ occurs for $1 \leq \Delta \leq 100$. The optimal choice of $\Delta$ is thus the smallest value which will provide decorrelation of the noise in which the signal is embedded. A fixed value of $\Delta = 2$ samples was used in the ROC performance analysis discussed below.

3) Sensitivity to Filter Length: The $P_D$ is plotted as a function of the ratio $L/K$ in Fig. 25 for a fixed $P_{FA}$, SNR, $\tau_a$, $\Delta$, and $T$. This figure illustrates that the ALE filter length necessary to provide maximum $P_D$ decreases as the signal bandwidth increases. For the cases shown here, the maximum $P_D$ occurs where $\alpha = \beta_1$ for all values of $\alpha$ considered. This indicates that performance of both the ALE output detector and the ALE weight detector is optimized by matching the ALE frequency resolution to the signal bandwidth. A general increase in the performance sensitivity to the choice of $L$ is indicated in Fig. 25 as the signal bandwidth increases. These results are qualitatively similar to those obtained in [60] for a linear FM signal. In both cases SNR enhancement decreases for $L > f_s/\alpha$ because the end weights of the filter primarily contribute additional misadjustment noise to the ALE output. This occurs because the correlation between $x(k)$ and $x(k - i - \Delta)$ becomes small as $i \rightarrow L$ for $L > f_s/\alpha$. Fig. 25 illustrates that the performance of the ALE output detector is superior to that of the ALE weight detector. Note, however, that the optimal filter length for both the weight and output detectors is equivalent to the value defined in [49] for maximum SNR enhancement, i.e.,

$$L_{opt} = f_s/\alpha$$  \hspace{1cm} (42)

E. Comparative Performance of ALE-Augmented and Conventional Detectors

Figure 26 indicates an overall degradation in detection performance of the adaptive detector as the signal bandwidth is increased. Fig. 26 plots the $P_D$ as a function of SNR for a fixed $P_{FA}$ of $10^{-4}$ for the ALE output detector. The solid curves represent the condition $\beta_1 = \beta_2 = 1$ Hz for the designated values of $\alpha$ between 0 and 4 Hz for the ALE output detector, while the dashed curves indicate performance for $L = L_{opt}$. Equivalent results for the ALE weight detector are shown in Fig. 27. The dashed curves in Fig. 27 again indicate
Fig. 25. Sensitivity to filter length for various signal bandwidths.

Fig. 26. Detection sensitivity to signal bandwidth for ALE output detector.

Fig. 27. Detection sensitivity to signal bandwidth for ALE weight detector.
that \( L = L_{opt} \). Figure 27 illustrates that for \( \alpha = 4 \) Hz, gains in detection performance of 2 dB at \( P_D = 50\% \) can be achieved simply by truncating the weight vector from 1024 to 256 to eliminate the algorithm noise which dominates the end weights of the filter. The performance of conventional detectors for the same input conditions is shown in Fig. 28.

The comparative performance of the ALE output detector, the ALE weight detector, and the conventional detector is plotted as a function of input signal bandwidth for a fixed \( P_D \) of 50\% and a fixed \( P_{FA} \) of \( 10^{-4} \) in Fig. 29. Figure 29 indicates that the performance of all three detectors degrades as the signal bandwidth is increased, but the performance of adaptive and conventional detectors remains nearly the same if \( L \) is chosen to ensure that \( \beta_1 = \alpha \). Figure 29 illustrates that the performance of the ALE weight detector and the ALE output detector is essentially the same for sinusoidal inputs, but that the performance of the ALE weight detector degrades more rapidly with increasing signal bandwidth, particularly when \( \beta_1 = \alpha \). A relative loss of 1.8 dB in performance between the ALE output detector and the conventional detector occurs as \( \alpha \) increases from 0 to 4 Hz for \( \beta_1 \) fixed at 1 Hz. The equivalent relative loss in performance between the ALE weight detector and the conventional detector is approximately 3.7 dB as \( \alpha \) increases from 0 to 4 Hz for \( \beta_1 = \beta_2 = 1 \) Hz. The slight improvement in both the ALE weight detector and the ALE output detector for \( \alpha = 0 \) arises from implementation noise in the DFT processor [141]-[142].

**F. Detection Performance in Nonstationary Noise**

As shown above, the presence of the ALE produces a radical change in the noise-only pdf of the unintegrated DFT output. This is also true for the integrated output case, as illustrated by Figs. 30 and 31. In either instance, both the mean and the mean-to-variance ratio of the noise-only pdf are significantly reduced for the ALE/DFT detector, as compared with the conventional DFT detector. For the detection of a signal in a stationary noise background of known power, this effect is of no consequence, since accurate
thresholds can be established for both detectors from a knowledge of the noise-only pdf. When the noise power is unknown, however, the threshold level must be estimated from the received data. In this case, any error in estimating the noise power spectral density will result in bias errors in the detection threshold. When the noise is nonstationary, a new estimate of the noise power spectral density must be made at each detection interval. When the signal duration is unknown a priori and postdection integration is employed to improve detection performance, the estimation                                  errors accumulate with each new coherent detector output that is added to form the integrated detector output.

The difficulty in providing accurate estimates of the noise power spectral density is further complicated when the signal frequency, bandwidth, and duration are unknown a priori and the background noise is a composite of narrow-band and broadband terms, which produce nonstationary and nonuniform fluctuations in the power spectral density over the frequency range of interest. A variety of noise spectrum equalization and power spectral density estimation algorithms have been employed in such cases [128]-[150]. In these applications, the reduced mean-to-variance ratio of the ALE/DFT detector may prove advantageous if alternate means of estimating appropriate time-varying thresholds either prove ineffective or impractical. This is illustrated by the following simulation results.

We consider a zero-mean, stationary white Gaussian process \( n(k) \) having variance \( \sigma^2 \). Let \( m(t) \) denote an independent process of the same type, having variance \( \sigma^2_m \), where

\[
t = \left( \frac{k}{T} \right) \left( \text{integer part of} \frac{k}{T} \right)
\]

and \( T \) is the number of input samples spanned by the total integration interval of the detector. Thus the process \( m(t) \) changes value every \( T \) samples of \( n(k) \). Let the input noise to the detector during a particular integration interval \( [T, (t + 1)T] \) be given by

\[
n'(k) = [1 + m(t)] n(k), \quad T \leq k \leq (t + 1)T.
\]

Over the infinite time interval, the noise \( n'(k) \) is a block-stationary, non-Gaussian process, for which a time averaged (over time) false-alarm threshold may be determined via simulation. However, with respect to each integration interval of length \( T \), the noise variance is a random variable \( [1 + m(t)]^2 \sigma^2 \), conditioned on \( m(t) \), and thus \( n'(k) \) is nonstationary insofar as the detector is concerned. The degree of nonstationarity is described by the ratio \( \sigma^2_m / \sigma^2 \).

Figure 32 shows a comparison of the degradations in \( P_d \) (for fixed SNR) as a function of \( \tau_a / T \) for the conventional DFT and ALE/DFT detectors with incoherent integration. In this simulation, the ALE weights were reset to zero at the beginning of each integration interval of length \( T \). Note that the conventional system degrades more rapidly, and at higher
SNR values, than the adaptive system. This is to be expected, since the higher mean-to-variance ratio of the conventional system's noise-only pdf indicates a greater sensitivity to changes in the input noise power. Thus as the degree of noise nonstationarity increases, the false-alarm threshold for the conventional system must be increased by a proportionately greater amount than for the adaptive system in order to maintain a given average $P_{fa}$.

**Conclusions**

It was shown that the adaptive prediction filter can process bandlimited data as either desired signal or undesired noise by proper selection of the prediction distance, $\Delta$. The optimal value of $\Delta$ for noise suppression lies in an interval between the correlation distance of the signal and the correlation distance of the noise. A much tighter bound on $\Delta$ is necessary to achieve optimal resolution of multiple frequency components in the input if the components are closely spaced. The optimal value of $\Delta$ for frequency resolution is defined in [56], [93].

It was shown that the output of an adaptive prediction filter can be accurately modeled by a sum of Wiener filter components and other components which arise from the weight misadjustment noise associated with real-time Wiener filtering. Expressions were derived for the relative magnitude of the Wiener filter components and the misadjustment noise components for the LMS implementation of the adaptive prediction filter for narrowband signals of bandwidth, $\alpha$, in WGN. It was shown that the Wiener filter terms dominate at high input SNR, but below a threshold SNR, the ratio of the Wiener filter terms and the misadjustment terms decreases nonlinearly with decreasing input SNR. It was shown that at low SNR, the threshold performance can be controlled by proper selection of the filter length, $L$, and the adaptive filter time constant. It was shown that the optimal value of $L$ at low SNR is equal to $1/\alpha$. Further increase in $L$ overresolves the narrowband signal components and increases filter misadjustment noise. It was shown that this optimal choice $L$ for signal enhancement also defines the optimal ROC detection performance bounds. It is further shown that the Wiener filter bandwidth significantly exceeds the input signal bandwidth at high SNR, but approaches the input signal bandwidth at low SNR. Consequently, parameter estimators based on the adaptive filter output will often provide superior performance to estimators based on the adaptive filter weights.

Two different detector implementations of the adaptive prediction filter were compared using ROC curves obtained by computer simulation. It was shown that due to the faster convergence of the prediction filter output and the improved correlation between the filter output and the input signal, the ROC performance of an adaptive detector based on the prediction filter output is superior to that of an adaptive detector based on the prediction filter weights. It was shown that the ROC performance of the adaptive detector implementation based on the prediction filter output will closely approach that of the ideal matched filter energy detector. These results illustrate that the signal enhancements which are evident above the threshold SNR are offset by changes in the output noise statistics. These factors trade off in a manner which ensures that the detect-
tion performances of the LMS implementation of the adaptive prediction filter closely approximates that of an optimal Wiener filter.

ACKNOWLEDGEMENT

I am grateful to Prof. Walter H. Ku for his encouragement and support and to Prof. Larry Milstein, Dr. J. T. Rickard, Prof. Bhaskar Rao, Kay Chew, and Dr. David Stein for their assistance and comments. I am also indebted to Drs. Ed Satoriis, Mark Shensa, J. T. Rickard, Mauro Dentino, and Candace Anderson for their contributions to this research, and to Debra LaPierre for her assistance in preparing the manuscript.

REFERENCES


James R. Zeidler (Senior Member, IEEE) received the B.A. degree in Physics from MacMurray College, Jacksonville, Ill in 1966, the M.S. degree in Physics from Michigan State University in 1968, and the Ph.D. degree in Physics from the University of Nebraska in 1972. He received a fellowship to the NASA Summer Institute in Space Physics at Columbia University, New York, N.Y. in 1966. In 1973 he was a postdoctoral research associate in the Department of Electrical Engineering at the University of Nebraska and a lecturer in Engineering Physics.

Since 1974 he has been a scientist at the Naval Ocean Systems Center, San Diego, C.A., where he has done research and development in adaptive signal processing, underwater acoustic communications, satellite communications, and in diamond electronic devices. From 1983 to 1984 he was a technical assistant in the Office of the Assistant Secretary of the Navy (Research, Engineering, and Systems), Washington, D.C. Since 1988 he has also been affiliated with the University of California at San Diego, where he is currently Adjunct Professor of Electrical and Computer Engineering. He has written numerous papers and has six patents in the areas of signal processing and solid state physics.