Comparative Tracking Performance of the LMS and RLS Algorithms for Chirped Narrowband Signal Recovery

Paul C. Wei, *Member, IEEE*, Jun Han, *Student Member, IEEE*, James R. Zeidler, *Fellow, IEEE*, and Walter H. Ku, *Member, IEEE*

***Abstract*—This paper studies the comparative tracking performance of the recursive least squares (RLS) and least mean square (LMS) algorithms for time-varying inputs, specifically for linearly chirped narrowband input signals in additive white Gaussian noise. It is shown that the structural differences in the implementation of the LMS and RLS weight updates produce regions where the LMS performance exceeds that of the RLS and other regions where the converse occurs. These regions are shown to be a function of the signal bandwidth and signal-to-noise ratio (SNR). LMS is shown to place a notch in the signal band of the mean lag filter, thus reducing the lag error and improving the tracking performance. For the chirped signal, it is shown that this produces smaller tracking error for small SNR. For high SNR, there is a region of signal bandwidth for which RLS will provide lower error than LMS, but even for these high SNR inputs, LMS always provides superior performance for very narrowband signals.**

***Index Terms*—Adaptive filters, autoregressive model, least mean square, recursive least squares, tracking.**

I. Introduction

THE tracking behavior of adaptive filtering algorithms is a fundamental issue in defining their performance in nonstationary operating environments. It has been established [1]–[3] that adaptive algorithms that exhibit good convergence properties in stationary environments do not necessarily provide good tracking performance in a nonstationary environment because the convergence behavior of an adaptive filter is a transient phenomenon, whereas the tracking behavior is a steady-state property.

There are two fundamentally different nonstationary scenarios that have been previously analyzed. The first involves a time-varying system where the cross-correlation vector between the input signal to the adaptive filter and the desired response is time varying; the second involves a nonstationary input to the adaptive filter. The first case occurs in system identification applications and is analyzed in [1], [2], and [4]. The latter case occurs in adaptive equalization, signal recovery, and other applications and has been analyzed in [3], [5], and [6] for the least mean square (LMS) and the recursive least squares (RLS) adaptive algorithms for the specific case of a deterministic chirped signal input to an adaptive prediction filter. This signal model provides a constant nonstationarity that is useful in separating the convergence and tracking behavior. The results show that LMS usually provides better tracking behavior than exponentially weighted RLS for linearly chirped sinusoids in additive white Gaussian noise (AWGN). This behavior is because LMS is model independent, whereas RLS must employ a model of the data correlation matrix (such as an exponential weighting), which may not match the characteristics of the input signals. It is further shown in [7] that superior tracking performance is obtained by RLS if the extended RLS (ERLS) algorithm with a time-varying state transition matrix is used. For chirped inputs, the ERLS implementation estimates the chirp rate from the data and uses it thereafter in the adaptive update equations [7]. Nevertheless the exponentially weighted RLS algorithm is frequently employed in practical systems due to its computational simplicity compared with ERLS and because, in many scenarios, an accurate state-space model is not available. Consequently, a comparison of the tracking performance of the LMS and the exponentially weighted RLS algorithm remains an important issue.

The tracking behavior of the RLS adaptive filter was extended to include stochastic narrowband input signals in [8]. The signal bandwidth is shown to be an important parameter in optimizing the RLS filter performance [8]. This paper extends the analysis in [3], [5], [6], and [8] to define the tracking performance of the LMS adaptive algorithm for narrowband stochastic input signals using the same signal model used in [8]. This will allow the comparative performance of the LMS and RLS algorithms to be defined for a broader class of input signals and to determine the effects of signal bandwidth on the tracking performance. This is shown to provide a number of interesting results for the adaptive signal recovery application.

The structural differences in the LMS and RLS weights updates are shown to produce regions where the LMS performance exceeds that of the RLS and vice versa. These regions are functions of the signal bandwidth and signal-to-noise ratio (SNR). LMS is shown to place a notch in the signal band of the mean lag filter, thus reducing the lag error and improving the tracking

Manuscript received April 6, 2001; revised March 4, 2002. This work was supported by the National Science Foundation Industry/University Cooperative Research Center on Ultra-High Speed Integrated Circuits and Systems (ICAS) at the University of California, San Diego. The associate editor coordinating the review of this paper and approving it for publication was Dr. Dennis R. Morgan.

P. C. Wei, J. Han, and W. H. Ku are with the Department of Electrical and Computer Engineering, University of California, San Diego, La Jolla, CA 92093-0407 USA.

J. R. Zeidler is with the Department of Electrical and Computer Engineering, University of California, San Diego, La Jolla, CA 92093-0407 USA, and with the Space and Naval Warfare Center, San Diego, CA 92152 USA (e-mail: zeidler@ece.ucsd.edu).

Publisher Item Identifier S 1053-587X(02)05640-4.

performance. The optimal adaptation constant for both LMS and RLS with chirped narrowband inputs is shown to yield a lag misadjustment error that is half the noise misadjustment error, which is identical to previous results for deterministic inputs [3], [5]. This is in contrast to the result obtained for systems with time-varying coefficients [4], where optimal tracking occurs when the lag and noise misadjustment are equal.

II. LMS Algorithm

In order to facilitate the steady-state tracking performance comparison between the LMS and RLS algorithms for a chirped narrowband signal in AWGN, the problem introduction, the chirped narrowband signal model, the optimum predictor, and notations used in this paper will be the same as described in [8], where the steady-state tracking performance of the RLS algorithm is studied and contrasted with the results in [3]. To complete the comparisons, we adopt the notational conventions in the previous studies [3], [5], [6], and [8].

The LMS adaptive one-step forward predictor is summarized by the following filter output and weight update equations. The filter output is given by

$$y_k = \bar{W}_{k-1}^T \bar{x}_k \tag{1}$$

and the weight update equation is

$$\bar{W}_k = \bar{W}_{k-1} + \mu \bar{x}_k^* (d_k - y_k). \tag{2}$$

The LMS algorithm uses the instantaneous estimate of the input correlation ($\bar{x}_k^* \bar{x}_k^T$) and cross correlation ($\bar{x}_k^* d_k$) to update the filter weights [2].

The LMS weight update equation is used to decompose the weight error into a mean and a fluctuation component. The difference between the LMS weights and the optimum Wiener weights is [5]

$$\bar{\Delta}_{k-1} = \bar{W}_{k-1} - \bar{W}_k^o. \tag{3}$$

The adaptive filter output misadjustment is defined as

$$\mathcal{M}_k = E\left|\bar{\Delta}_{k-1}^T \bar{x}_k\right|^2. \tag{4}$$

The LMS filter weight update equation can be rewritten as

$$\bar{W}_k = [\mathbf{I} - \mu \bar{x}_k^* \bar{x}_k^T] \bar{W}_{k-1} + \mu \bar{x}_k^* x_k \tag{5}$$

where the desired signal d_k in this case is equal to x_k. Subtracting $\bar{W}_{k+1}^o$ from both sides yields

$$\bar{\Delta}_k = (\mathbf{I} - \mu \bar{x}_k^* \bar{x}_k^T) \bar{\Delta}_{k-1} - \bar{T}_k + \mu \bar{x}_k^* e_k^o \tag{6}$$

where $\bar{T}_k$ is the change in the optimum weight vector at successive iteration

$$\bar{T}_k = \bar{W}_{k+1}^o - \bar{W}_k^o = \mathbf{V}^{k+1} \mathbf{\Lambda} (\bar{W}^o \odot \bar{D}) \tag{7}$$

$\mathbf{\Lambda}$ is a diagonal matrix given by

$$\mathbf{\Lambda} = \mathbf{I} - \mathbf{V}^* \tag{8}$$

and $\mathbf{V}$ is the chirp matrix, $\bar{D}$ is the signal direction vector as defined in [8]

$$\mathbf{V} = \text{diag}\left(\Psi, \Psi^2, \ldots, \Psi^M\right) \tag{9}$$

$$\bar{D} = \left[\Omega \Psi^{-1/2}, \Omega^2 \Psi^{-2^2/2}, \ldots, \Omega^M \Psi^{-M^2/2}\right]^T \tag{10}$$

where $\Psi = e^{j\psi}$, ψ is the chirp rate, and $\Omega = e^{j\omega_0}$, ω_0 defines the initial center frequency of the spectrum. The predictor error when using the optimum weights is

$$e_k^o = x_k - \left(\bar{W}_k^o\right)^T \bar{x}_k. \tag{11}$$

Employing the same assumptions utilized in [5] that $\bar{x}_k$ is a sequence of independent vectors, it follows that the optimum prediction error e_k^o is uncorrelated with the input, i.e.,

$$E\left\{\left[x_k - \left(\bar{W}_k^o\right)^T \bar{x}_k\right] \bar{x}_k^*\right\} = 0 \tag{12}$$

and thus, the noise and lag weight errors are independent. Since the weight error update equation is linear, it can be decomposed into the constituent terms

$$\bar{\Delta}_k = \bar{\Delta}_k^n + \bar{\Delta}_k^l \tag{13}$$

consisting of a noise weight error vector and a lag weight error vector with the following update equations:

$$\bar{\Delta}_k^n = \left[\mathbf{I} - \mu \bar{x}_k^* \bar{x}_k^T\right] \bar{\Delta}_{k-1}^n + \mu \bar{x}_k^* e_k^o \tag{14}$$

$$\bar{\Delta}_k^l = \left[\mathbf{I} - \mu \bar{x}_k^* \bar{x}_k^T\right] \bar{\Delta}_{k-1}^l - \bar{T}_k. \tag{15}$$

Similar to the analysis for the RLS algorithm [8], the lag weight error can be further decomposed into a mean component $E[\bar{\Delta}_k^l]$ and a fluctuation component $\tilde{\Delta}_k^l = \bar{\Delta}_k^l - E[\bar{\Delta}_k^l]$. It will be shown that under the "slow adaptation" condition employed in [3] and [5], the misadjustment from the lag fluctuation $\tilde{\Delta}_k^l$ is small compared with the mean lag component. Thus, the filter can be viewed as three independent filters, consisting of a Wiener filter, noise filter, and lag filter component. The output misadjustment therefore consists of the noise, mean lag, and lag fluctuation misadjustment

$$\mathcal{M}_k = E\left[\left(\bar{\Delta}_{k-1}\right)^H \mathbf{\Phi}_k^x \bar{\Delta}_{k-1}\right] = \mathcal{M}_k^n + \left(\mathcal{M}_k^l + \tilde{\mathcal{M}}_k^l\right) \tag{16}$$

where $\mathbf{\Phi}_k^x$ is the input autocorrelation matrix, which can be written as [8]

$$\mathbf{\Phi}_k^x = E\left[\bar{x}_k^* \bar{x}_k^T\right] = P_n \mathbf{V}^k \mathcal{D} \mathbf{V}^{*k} \tag{17}$$

and

$$\mathcal{D} = \left[\mathbf{I} + \rho \mathbf{R} \odot \left(\bar{D} \bar{D}^H\right)\right] \tag{18}$$

where $\rho = P_s/P_n$ is the input SNR. To make further comparisons, we will now define a normalized adaptation constant for the LMS to remove the units of power

$$\nu = P_n(1+\rho)\mu. \tag{19}$$

In the following derivation of the misadjustment, we will assume the same "slow chirping" and "slow adaptation" assumption as used in [3] and [5], i.e.,

$$\text{SC}: M\psi \ll \mu \tag{20}$$

$$\text{SA}: M\nu \ll 2. \tag{21}$$

A. Noise Misadjustment

Since the noise and lag components are uncorrelated, they can be evaluated separately. The noise weight error update in (14) is the same as the stationary results [2]. The steady-state ($k \to \infty$) noise misadjustment is

$$\mathcal{M}^n = \frac{M\nu}{2 - M\nu}\xi_0 \tag{22}$$

when

$$0 < \mu < \frac{2}{\lambda_{\max}} \tag{23}$$

to assure the convergence in the mean squared sense, where $\lambda_{\max}$ is the largest of all the eigenvalues of the input correlation matrix $\mathbf{\Phi}_k^x$, $\lambda_i, i = 1, 2, \ldots, M$. By noting that $\sum_{i=1}^{M} \lambda_i = \text{Tr}(\mathbf{\Phi}_k^x) = MP_n(1+\rho)$, it can be seen that this condition is easily satisfied through the slow adaptation assumption, i.e., $\mu\text{Tr}(\mathbf{\Phi}_k^x) = \mu MP_n(1+\rho) = M\nu \ll 2$.

B. Mean Lag Misadjustment

The lag misadjustment is composed of two parts: the mean lag and the lag fluctuation. The mean lag weights $E[\bar{\Delta}_k^l]$ are given by the recursion

$$\begin{aligned} E\left[\bar{\Delta}_k^l\right] =& \left[\mathbf{I} - \mu\mathbf{\Phi}_k^x\right] E\left[\bar{\Delta}_{k-1}^l\right] - \bar{T}_k \\ =& \mathbf{V}^k\left(\mathbf{I} - \mu P_n\mathcal{D}\right)\mathbf{V}^{-k} E\left[\bar{\Delta}_{k-1}^l\right] \\ & - \mathbf{V}^k(\mathbf{V} - \mathbf{I})\left(\bar{W}^o \odot \bar{D}\right). \end{aligned} \tag{24}$$

Letting $\bar{Q}_k = \mathbf{V}^{-(k+1)} E[\bar{\Delta}_k^l]$, the dependence of $E[\bar{\Delta}_k^l]$ on time k can be removed, and the above simplifies to

$$\bar{Q}_k = \mathbf{V}^*\left(\mathbf{I} - \mu P_n\mathcal{D}\right)\bar{Q}_{k-1} - \left(\mathbf{I} - \mathbf{V}^*\right)\left(\bar{W}^o \odot \bar{D}\right). \tag{25}$$

This is a first-order difference equation with constant coefficients. When $(\mathbf{I} - \mu P_n\mathcal{D})^{-1}$ exists, $\bar{Q}_k$ converges to a steady-state value given by

$$\bar{Q} = -\left(\mathbf{\Lambda} + \mu P_n\mathcal{D}\right)^{-1}\mathbf{\Lambda}\left(\bar{W}^o \odot \bar{D}\right). \tag{26}$$

Using the solution for $\bar{Q}_k$ and solving for $E[\bar{\Delta}_k^l]$ gives the mean lag weight vector in the steady state as

$$\begin{aligned} E\left[\bar{\Delta}_k^l\right] =& -\mathbf{V}^{k+1}\left(\mathbf{\Lambda} + \mu P_n\mathcal{D}\right)^{-1}\mathbf{\Lambda}\left(\bar{W}^o \odot \bar{D}\right) \\ \approx& -\frac{1}{\mu P_n}\mathbf{V}^{k+1}\mathcal{D}^{-1}\mathbf{\Lambda}\left(\bar{W}^o \odot \bar{D}\right). \quad \text{(SC)} \end{aligned} \tag{27}$$

The lag misadjustment error is given by

$$\mathcal{M}_k^l = \text{Tr}\left(E\left[\bar{\Delta}_{k-1}^l\right]^H \mathbf{\Phi}_k^x E\left[\bar{\Delta}_{k-1}^l\right]\right) = P_n\bar{Q}_{k-1}^H\mathcal{D}\bar{Q}_{k-1}. \tag{28}$$

At steady state ($k \to \infty$), for $\psi \ll \mu P_n(1+\rho)$

$$\begin{aligned} \mathcal{M}^l \approx& P_n\left(\bar{W}^o \odot \bar{D}\right)^H \mathbf{\Lambda}^*\left(\mu P_n\mathcal{D}\right)^{-1}\mathcal{D}\left(\mu P_n\mathcal{D}\right)^{-1} \\ \cdot& \mathbf{\Lambda}\left(\bar{W}^o \odot \bar{D}\right) = \frac{\kappa_{\text{LMS}}}{\nu^2}\psi^2 \end{aligned} \tag{29}$$

where

$$\kappa_{\text{LMS}} = \left[P_n(1+\rho)\right]^2 \frac{\left(\mathbf{\Lambda}\bar{W}^o\right)^H\left(\mathbf{\Phi}^x\right)^{-1}\left(\mathbf{\Lambda}\bar{W}^o\right)}{\psi^2} \tag{30}$$

is the *normalized lag misadjustment* of LMS for a chirped process. In (30), $\mathbf{\Phi}^x$ is the autocorrelation matrix of the corresponding stationary input process [8]. Note that κ_{LMS} has the units of power and is *independent* of the chirp rate under the slow chirp condition since the diagonal matrix $\mathbf{\Lambda}$ is linear with respect to ψ with elements $\mathbf{\Lambda}_{l,l} \approx j\psi l, l \in [1, 2, \ldots M]$.

C. Lag Fluctuation Misadjustment

In this section, it is shown that the misadjustment caused by the lag fluctuation is small compared with the mean lag misadjustment. The lag fluctuation misadjustment is given by

$$\tilde{\mathcal{M}}_k^l = \text{Tr}\left(\mathbf{\Phi}_k^x Z_k\right) \tag{31}$$

where

$$\mathbf{Z}_k = E\left[\tilde{\Delta}_{k-1}^l\left(\tilde{\Delta}_{k-1}^l\right)^H\right] \tag{32}$$

is the lag fluctuation weight correlation matrix, and the lag fluctuation weight is given by

$$\begin{aligned} \tilde{\Delta}_k^l =& \bar{\Delta}_k^l - E\left[\bar{\Delta}_k^l\right] \\ =& \left(\mathbf{I} - \mu\bar{x}_k^*\bar{x}_k^T\right)\tilde{\Delta}_{k-1}^l + \mu\left(\mathbf{\Phi}_k^x - \bar{x}_k^*\bar{x}_k^T\right)E\left[\bar{\Delta}_{k-1}^l\right]. \end{aligned} \tag{33}$$

For the linear chirp model defined above, we can assume that $\mathbf{Z}_k$ takes the form $\mathbf{Z}_k = \mathbf{V}^k\mathbf{B}_k\mathbf{V}^{*k}$. This allows the lag fluctuation misadjustment to be simplified to

$$\tilde{\mathcal{M}}_k^l = P_n\text{Tr}\left\{\mathcal{D}\mathbf{B}_k\right\}. \tag{34}$$

Using (33) and the identity in the Appendix, the lag fluctuation weight correlation $\mathbf{Z}_k$ follows the update equation

$$\begin{aligned} \mathbf{Z}_{k+1} =& \mathbf{Z}_k - \mu\left(\mathbf{\Phi}_k^x\mathbf{Z}_k + \mathbf{Z}_k\mathbf{\Phi}_k^x\right) \\ & + \mu^2\left\{\mathbf{\Phi}_k^x\mathbf{Z}_k\mathbf{\Phi}_k^x + \mathbf{\Phi}_k^x\text{Tr}\left[\mathbf{\Phi}_k^x\mathbf{Z}_k\right]\right\} \\ & + \mu^2 E\left[\bar{\Delta}_{k-1}^l\right]^H\mathbf{\Phi}_k^x E\left[\bar{\Delta}_{k-1}^l\right]\mathbf{\Phi}_k^x. \end{aligned} \tag{35}$$

Substituting the form $\mathbf{Z}_k = \mathbf{V}^k\mathbf{B}_k\mathbf{V}^{*k}$, then (35) becomes

$$\begin{aligned} \mathbf{V}\mathbf{B}_{k+1}\mathbf{V}^* =& \mathbf{B}_k - \mu P_n\left(\mathbf{B}_k\mathcal{D} + \mathcal{D}\mathbf{B}_k\right) \\ & + \mu^2 P_n^2\left\{\mathcal{D}\mathbf{B}_k\mathcal{D} + \mathcal{D}\text{Tr}\left[\mathcal{D}\mathbf{B}_k\right]\right\} + \mu^2 P_n\mathcal{M}_k^l\mathcal{D}. \end{aligned} \tag{36}$$

At steady state ($k \to \infty$), $\mathbf{B}_k \to \mathbf{B}$. A bound of $\text{Tr}(\mathcal{D}\mathbf{B})$ can be obtained by taking the trace and then applying the two-norm properties of the correlation matrices to give

$$\begin{aligned} \left\{2\mu P_n - \mu^2 P_n^2\text{Tr}[\mathcal{D}]\right\}\text{Tr}[\mathcal{D}\mathbf{B}] =& \mu^2 P_n^2\text{Tr}[\mathcal{D}\mathbf{B}\mathcal{D}] \\ & + \mu^2 P_n\text{Tr}[\mathcal{D}]\mathcal{M}_k^l. \end{aligned} \tag{37}$$

TABLE I
SUMMARY OF OPTIMUM NORMALIZED ADAPTATION CONSTANT AND MISADJUSTMENT

Parameter	RLS Algorithm	LMS Algorithm
Normalized adaptation constant (β, ν)	$\beta = 1-\lambda$	$\nu = P_n(1+\rho)\mu$
$\kappa_{\mathrm{RLS}}, \kappa_{\mathrm{LMS}}$	$(\mathbf{\Lambda}\bar{W}_o)^H \mathbf{\Phi}^x (\mathbf{\Lambda}\bar{W}_o)/\psi^2$	$[P_n(1+\rho)]^2 (\mathbf{\Lambda}\bar{W}_o)^H (\mathbf{\Phi}^x)^{-1} (\mathbf{\Lambda}\bar{W}_o)/\psi^2$
Misadjustment $\mathcal{M}$	$\frac{\xi_o M}{2}\beta + \frac{\psi^2}{\beta^2}\kappa_{\mathrm{RLS}}$	$\frac{\xi_o M}{2}\nu + \frac{\psi^2}{\nu^2}\kappa_{\mathrm{LMS}}$
Optimal normalized adaptation constant ($\beta_{\mathrm{opt}}, \nu_{\mathrm{opt}}$)	$\beta_{\mathrm{opt}} = \left(\frac{4\psi^2}{M}\frac{\kappa_{\mathrm{RLS}}}{\xi_o}\right)^{1/3}$	$\nu_{\mathrm{opt}} = \left(\frac{4\psi^2}{M}\frac{\kappa_{\mathrm{LMS}}}{\xi_o}\right)^{1/3}$
Optimal misadjustment $\mathcal{M}_{\min}$	$\frac{3}{4}\xi_o M \beta_{\mathrm{opt}}$	$\frac{3}{4}\xi_o M \nu_{\mathrm{opt}}$

and

$$\begin{aligned}\{2-\mu P_n \mathrm{Tr}[\mathcal{D}]\}\,\mathrm{Tr}[\mathcal{D}\mathbf{B}] &= \mu P_n \mathrm{Tr}[\mathcal{D}\mathbf{B}\mathcal{D}] + \mu \mathcal{M}_k^l \mathrm{Tr}[\mathcal{D}] \\ &\le \mu P_n \|\mathcal{D}\| \mathrm{Tr}[\mathbf{B}\mathcal{D}] \\ &\quad + \mu \mathrm{Tr}[\mathcal{D}]\mathcal{M}_k^l. \end{aligned} \tag{38}$$

Thus

$$\begin{aligned}\tilde{\mathcal{M}}^l = \tilde{\mathcal{M}}_{k\to\infty}^l = P_n \mathrm{Tr}[\mathcal{D}\mathbf{B}] &\le \frac{\mu P_n \mathrm{Tr}[\mathcal{D}]\mathcal{M}^l}{2-\mu P_n(\mathrm{Tr}[\mathcal{D}] + \|\mathcal{D}\|)} \\ &\le \frac{\mu P_n \mathrm{Tr}[\mathcal{D}]\mathcal{M}^l}{2-2\mu P_n \mathrm{Tr}[\mathcal{D}]}.\end{aligned}$$

The ratio of steady-state lag fluctuation misadjustment versus mean lag fluctuation is thus

$$\frac{\tilde{\mathcal{M}}^l}{\mathcal{M}^l} \le \frac{\mu P_n \mathrm{Tr}[\mathcal{D}]}{2-2\mu P_n \mathrm{Tr}[\mathcal{D}]}. \tag{39}$$

By the "slow adaptation" assumption (21), i.e., $\mu P_n \mathrm{Tr}[\mathcal{D}] \ll 2$, the above ratio is much smaller than unity. Thus, the lag fluctuation can be neglected. This is similar in form to the solution for the chirped sinusoid problem [5].

D. Optimum Normalized Adaptation Constant

Using the previous results for the noise and lag misadjustment, the normalized adaptation constant ν defined in (19) can be optimized to minimize the filter misadjustment. The total misadjustment is given by

$$\mathcal{M} = \frac{M\nu}{2-M\nu}\xi_o + \frac{\psi^2}{\nu^2}\kappa_{\mathrm{LMS}}. \tag{40}$$

When $\nu \ll 1$, the misadjustment has the same form as the RLS algorithm defined in [8]. Solving for the minimum with respect to ν, the optimum normalized adaptation constant is

$$\nu_{\mathrm{opt}} = \left(\frac{4\psi^2\kappa_{\mathrm{LMS}}}{M\xi_0}\right)^{1/3}. \tag{41}$$

The optimum misadjustment is

$$\mathcal{M}_{\min} = \frac{3}{4}M\xi_o\nu_{\mathrm{opt}}. \tag{42}$$

This has the same form as the RLS result derived in [8]. In both cases, the lag misadjustment is equal to half the noise misadjustment. The main results obtained for the optimal normalized adaptation constants and misadjustment noise for the LMS algorithm obtained above and the RLS algorithm in [8] are summarized in Table I. It is evident from Table I that the main difference in performance between the LMS and RLS algorithms will be defined by the differences in κ_{LMS} and κ_{RLS}. These terms define the properties of the weight update structure of each algorithm. The performance differences that result will be evaluated later for the chirped AR1 signal model.

III. COMPARATIVE TRACKING PERFORMANCE OF THE LMS AND RLS ALGORITHMS FOR A CHIRPED AR1 SIGNAL IN NOISE

To illustrate the performance differences between the LMS and RLS algorithms for a chirped nonzero bandwidth signal in AWGN, the results obtained previously are applied to a chirped AR1 process embedded in AWGN [8]. The AR1 process can be used to model many narrowband signals such as binary phase-shift keying (BPSK), as discussed in [8] and [9].

As shown in [8] for RLS and (6) for LMS, in the ALE configuration, the weight error vector of the LMS and RLS algorithms is given by

$$\bar{\Delta}_k = \begin{cases} \left(\mathbf{I} - \mu\bar{x}_k^*\bar{x}_k^T\right)\bar{\Delta}_{k-1} + \mu\bar{x}_k^* e_k^o - \bar{T}_k, & \text{LMS} \\ \lambda\mathbf{R}_k^{-1}\mathbf{R}_{k-1}\bar{\Delta}_{k-1} + \mathbf{R}_k^{-1}\bar{x}_k^* e_k^o - \bar{T}_k, & \text{RLS} \end{cases} \tag{43}$$

where $\mathbf{R}_k$ is the RLS estimate of the input correlation matrix using an exponentially weighted estimate with weighting parameter $\lambda = 1-\beta$. The first term summarizes the form of the update of the algorithm. The second term gives the adaptation noise error. The last term gives the lag error. The major difference is that the RLS algorithm decorrelates the input using $\mathbf{R}_k^{-1}$. For the LMS algorithm, however, the weight error update uses successive updates to subtract components of the signal from the weight error. In the RLS algorithm, assuming that $\mathbf{R}_k \approx \mathbf{R}_{k-1}$,

the previous weight is merely scaled. The weight error update (43) is general and not specific to a particular signal model.

The fundamental difference between the LMS and RLS algorithms with respect to their tracking performance can be described by evaluating the mean weight error vectors of the two algorithms. This is obtained by examining the expected value of (43). Using the definition of the predictor error in (11), the mean of e_k^o is zero. For the LMS algorithm, the steady-state solution to the first-order stochastic difference equation is $E[\bar{\Delta}_k] \approx -(1/\mu)E[\bar{x}_k^* \bar{x}_k^T]^{-1}\bar{T}_k$. The RLS solution was obtained in [8], where it was shown that for the slow chirp assumption, $\mathbf{R}_k^{-1}\mathbf{R}_{k-1} \approx \mathbf{I}$, and the weight difference of the RLS algorithm is thus $E[\bar{\Delta}_k] \approx -(1/\beta)\bar{T}_k$. The optimum weight difference vector $\bar{T}_k$ is defined by (7); thus, the mean weight error vectors for the LMS and RLS algorithms are given by

$$E\left[\bar{\Delta}_k\right] \approx \begin{cases} -\frac{1}{\mu}\mathbf{V}^{k+1}\left(P_n\mathcal{D}\right)^{-1}\mathbf{\Lambda}\left(\bar{W}^o \odot \bar{D}\right), & \text{LMS} \\ -\frac{1}{\beta}\mathbf{V}^{k+1}\mathbf{\Lambda}\left(\bar{W}^o \odot \bar{D}\right), & \text{RLS.} \end{cases} \tag{44}$$

The only major difference is the multiplication by $\mathcal{D}^{-1}$ in the LMS mean lag weight. The effect of this on the lag misadjustment is illustrated by decomposing $\mathcal{D}$ as

$$\mathcal{D} = \mathbf{I} + \rho\mathbf{R}_\omega \tag{45}$$

where $\mathbf{R}_\omega = \mathbf{R} \odot (\bar{D}\bar{D}^H)$. Under the slow chirp condition ($M\psi \ll v$), $\mathbf{R}_\omega$ is essentially the normalized correlation matrix of a stationary signal with initial frequency ω.

To illustrate the effects of $\mathcal{D}^{-1}$ on the filter performance, we can expand $\mathcal{D}$ using its eigencomponents

$$\mathcal{D} = \mathbf{I} + \rho\sum_{i=1}^{M}\sigma_i^2\bar{u}_i\bar{u}_i^H \tag{46}$$

where $\{\sigma_i^2, u_i\} i = 1 \ldots M$ are eigenvalues and associated eigenvectors of the signal correlation matrix $\mathbf{R}_\omega$. Since the eigenvectors $\{\bar{u}_i\}$ are orthonormal

$$\mathcal{D}^{-1} = \mathbf{I} - \sum_{i=1}^{M}\frac{\rho\sigma_i^2}{1+\rho\sigma_i^2}\bar{u}_i\bar{u}_i^H. \tag{47}$$

From (47), it is easy to see that the $\mathcal{D}^{-1}$ essentially subtracts components of the signal from the output. This has the effect of putting a notch in the signal band of the mean lag filter, as will be illustrated later.

The differences exhibited by the mean lag in the LMS and RLS algorithm defined in (44) are best illustrated by a plot of the adaptive filter transfer function. Fig. 1 and [8, Fig. 8], respectively, plot the components of the filter transfer function at three normalized adaptation constants for the LMS and RLS algorithms. The normalized adaptation constants are chosen to compare the effect of the notch at the optimal normalized adaptation constant ν_{opt} defined by (41) and normalized adaptation constant of $2\nu_{\text{opt}}$ and $\nu_{\text{opt}}/2$. Both the mean filter transfer function and the lag filter transfer function are shown. Note that the difference in the lag filter transfer function of the LMS and RLS

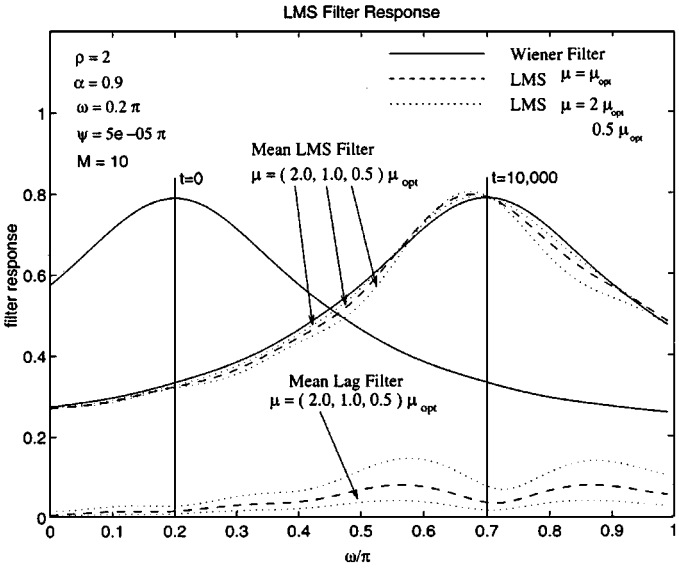

Fig. 1. LMS filter transfer function over time for a chirped AR1 process.

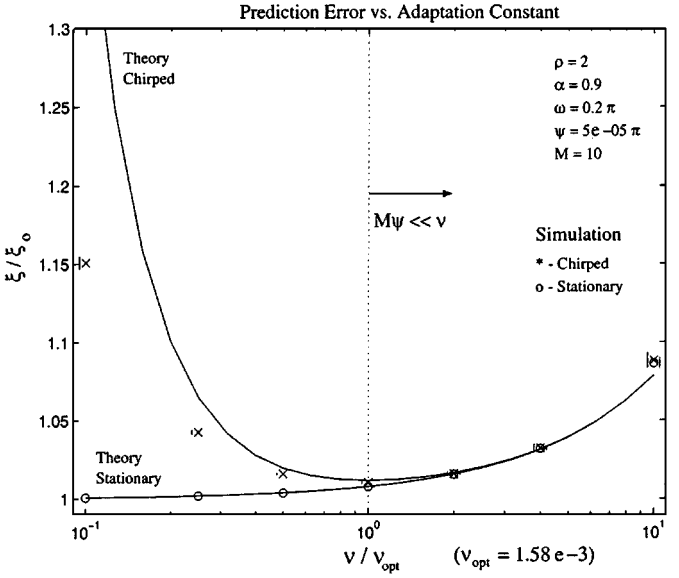

Fig. 2. Theoretical and experimental output misadjustment of the LMS algorithms for a chirped and a stationary AR1 input, plotted as a ratio to the Wiener MSE.

algorithms results from the effects of $\mathcal{D}^{-1}$ shown in (47). This produces a notching effect around the signal spectrum for the LMS algorithm that is not present in the RLS algorithm.

Fig. 2 and [8, Fig. 9], respectively, plot the output misadjustment of the LMS and RLS algorithms above their respective optimum adaptation constant with parameter values $\rho = 2$, $\alpha = 0.9$, $\omega = 0.2\pi$, $\psi = 5 \times 10^{-5}\pi$, and $M = 10$ for both stationary and chirped input signals. The vertical lines indicate the standard deviation ($\pm\sigma$) about the simulated results. It can be seen that near the optimum, the experimental results agree closely with the analytical results. When $\nu \geq \nu_{\text{opt}}$, the filter tracks the chirped signal with negligible lag error, and the error output is predominately noise misadjustment. Consequently, the output misadjustments for chirped and stationary AR1 inputs

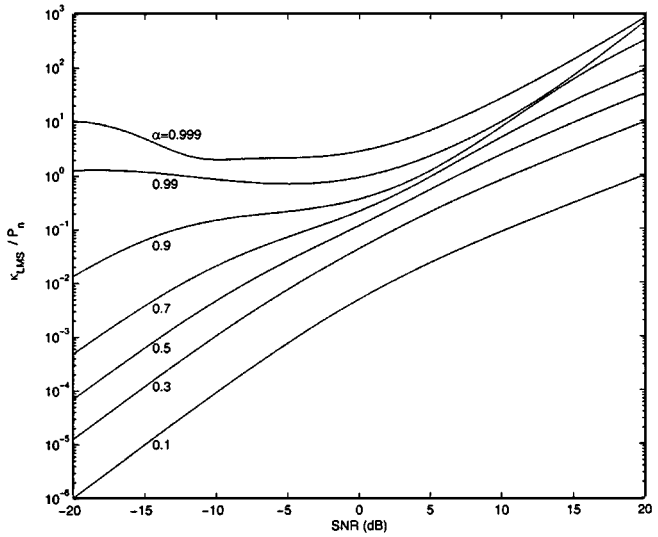

Fig. 3. Normalized asymptotic lag misadjustment $\lim_{M\to\infty}\kappa_{\mathrm{LMS}}/P_n$ versus SNR for a chirped AR1 signal with correlation α embedded in white noise. Results are shown for α between 0.1 and 0.99.

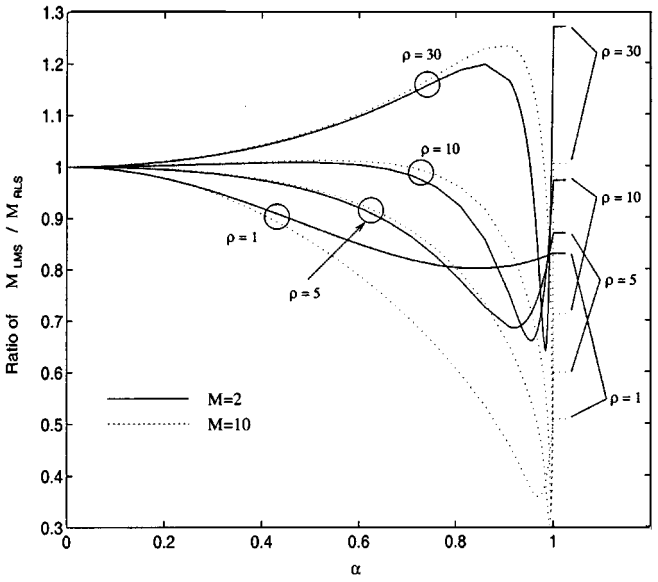

Fig. 4. Ratio of the LMS excess misadjustment vs. RLS excess misadjustment.

are almost identical. When $\nu \leq \nu_{\mathrm{opt}}$, the LMS does not track the chirped signal well, and the output error caused by the lag misadjustment dominates the performance. As a consequence, the difference between output error for the chirped and stationary AR1 inputs increases exponentially as $\nu/\nu_{\mathrm{opt}} \to 0$. The analytical results for the region where the slow chirp assumption is not valid are obtained numerically. Note that the simulation results indicate less misadjustment than the analysis predicts. These trends were also observed for the RLS algorithm in [8].

Additional comparisons between the analytical and simulation results are provided in [9]. The comparative misadjustment of the two algorithms can be plotted by substituting the optimum normalized adaptation constants from Table I into the expression for $\mathcal{M}_{\mathrm{opt}}$ and defining the ratio of the optimal output misadjustment for the LMS and RLS algorithms

$$\frac{(\mathcal{M}_{\min})_{\mathrm{LMS}}}{(\mathcal{M}_{\min})_{\mathrm{RLS}}} = \left(\frac{\kappa_{\mathrm{LMS}}}{\kappa_{\mathrm{RLS}}}\right)^{1/3} = \left(\frac{\left(\mathbf{\Lambda}\bar{W}^o\right)^H \mathcal{D}^{-1}\left(\mathbf{\Lambda}\bar{W}^o\right)}{\left(\mathbf{\Lambda}\bar{W}^o\right)^H \mathcal{D}\left(\mathbf{\Lambda}\bar{W}^o\right)}\right)^{1/3}. \quad (48)$$

Fig. 3 and [8, Fig. 3], respectively, plot the asymptotic normalized lag misadjustment parameter κ of the LMS and RLS algorithms normalized by the noise power (P_n) as the filter length $M \to \infty$. The normalized lag misadjustment for the two algorithms is equivalent for α between 0.1 and 0.7, but for α between 0.9 and 0.999, the differences in performance become apparent. Note from [8, Fig. 3] that κ_{RLS} increases monotonically as a function of SNR, whereas from Fig. 3, κ_{LMS} exhibits a nonmonotonic SNR response for narrowband signals with $\alpha \geq 0.9$.

Fig. 4 plots the ratio of the LMS misadjustment versus the RLS misadjustment for an AR1 process with $M = 2$ and 10 at various values of α and ρ, leading to the following observations.

1) As the signal bandwidth increases ($\alpha \to 0$), the ratio of the misadjustment approaches 1, i.e., in the limit where the signal and noise have the same bandwidth, both algorithms have the same performance.
2) For small SNR (approximately $\rho < 10$), the LMS algorithm always has a lower misadjustment. As the input SNR increases, however, the RLS algorithm has less error for a wide range of signal bandwidths. In all cases, however, there is a range of signal bandwidth where LMS outperforms RLS as $\alpha \to 1$. Note also that the range of α, where LMS outperforms RLS, becomes smaller as the SNR is increased. This property is explained by the differences in the ratios of $\mathcal{D}$ and $\mathcal{D}^{-1}$ in (46) and (47) as a function of ρ and α. This behavior is confirmed through simulations in [9].
3) The improvements for LMS become significant as $\alpha \to 1$. There is a discontinuity in the analysis for $\alpha = 1$, and the results obtained for a deterministic analysis under this condition are indicated by the bars at the $\alpha = 1$ axis for $\rho = 30$, 10, 5, and 1 for the cases of $M = 2$ and $M = 10$.

IV. Conclusions

The results show that the comparative tracking performance of the RLS and LMS algorithms for time-varying inputs such as the linear chirp in AWGN is highly dependent on the input signal bandwidth and signal-to-noise ratio ρ. Although both algorithms converge to the Wiener solution for stationary inputs, their performance in a nonstationary input environment can be traced back to their update structures. In this paper, the update structures of the LMS and RLS algorithms are examined. It is shown that the update of the LMS algorithm inherently subtracts signal components from the lag misadjustment. For the chirped signal, it is shown that this produces smaller tracking error for small SNR. In the chirped AR1 signal case, the LMS always has smaller tracking error when $\rho \leq 10$ dB. For $\rho \geq 10$ dB, it was shown that there is a region of signal bandwidths where RLS has superior performance, but even for these high SNR inputs, LMS has superior performance for very narrowband signals ($\alpha \to 1$). The range of bandwidths for which LMS out-

performs RLS was shown to be a function of ρ. It was further shown that the optimal performance for both the LMS and RLS algorithms was achieved by the use of an adaptation constant that provided a lag misadjustment of half the noise misadjustment for a stationary signal of the same properties. For adaptation constants greater than this optimal value, the output misadjustment error has a negligible component due to lag error and is dominated by the noise misadjustment term. Consequently, the performance of the adaptive filter for an adaptation constant greater than the optimum is the same for stationary and chirped inputs for both the LMS and RLS algorithms. For an adaptation constant less than the optimum, the lag misadjustment begins to dominate, and there is a significant difference in performance for the stationary and chirped inputs.

APPENDIX

In this Appendix, the quantities $E[\bar{x}_k^*\bar{x}_k^T\tilde{\Delta}_{k-1}^{\boldsymbol{l}}(\tilde{\Delta}_{k-1}^{\boldsymbol{l}})^H \bar{x}_k^*\bar{x}_k^T]$, $E\{\bar{x}_k^*\bar{x}_k^T E[\bar{\Delta}_{k-1}^{\boldsymbol{l}}]E[\bar{\Delta}_{k-1}^{\boldsymbol{l}}]^H\bar{x}_k^*\bar{x}_k^T\}$ are evaluated for the derivation from (33) to (35).

The following identity is used in the evaluation: For zero-mean Gaussian random vector $\bar{x}$ and vector $\bar{y}$, which can be either deterministic or random but, in general, uncorrelated with $\bar{x}$

$$E\left[\bar{x}^*\bar{x}^T\bar{y}\bar{y}^H\bar{x}^*\bar{x}^T\right] = \mathbf{R}_x\mathbf{R}_y\mathbf{R}_x + \mathbf{R}_x\mathrm{Tr}\left(\mathbf{R}_x\mathbf{R}_y\right)$$

where $\mathbf{R}_x = E[\bar{x}^*\bar{x}^T]$, $\mathbf{R}_y = E[\bar{y}\bar{y}^H]$ for random $\bar{y}$, $\mathbf{R}_y = \bar{y}\bar{y}^H$ for deterministic $\bar{y}$.

A. Proof

$$\begin{aligned}
\mathbf{U} &\triangleq \bar{x}^*\bar{x}^T\bar{y}\bar{y}^H\bar{x}^*\bar{x}^T\\
\mathbf{U}_{ij} &= \sum_{k=1}^{M}\sum_{l=1}^{M}\left(\bar{x}^*\bar{x}^T\right)_{ik}\left(\bar{y}\bar{y}^H\right)_{kl}\left(\bar{x}^*\bar{x}^T\right)_{lj}\\
E\left[\mathbf{U}_{ij}\right] &\triangleq \sum_{k=1}^{M}\sum_{l=1}^{M}E\left[\left(\bar{y}\bar{y}^H\right)_{kl}\right]E\left[\left(\bar{x}\right)_i^*\left(\bar{x}\right)_k\left(\bar{x}\right)_l^*\left(\bar{x}\right)_j\right]\\
&= \sum_{k=1}^{M}\sum_{l=1}^{M}E\left[\left(\bar{y}\bar{y}^H\right)_{kl}\right]\\
&\qquad\left\{E\left[\left(\bar{x}\right)_i^*\left(\bar{x}\right)_k\right]E\left[\left(\bar{x}\right)_l^*\left(\bar{x}\right)_j\right]\right.\\
&\qquad\left.+E\left[\left(\bar{x}\right)_i^*\left(\bar{x}\right)_j\right]E\left[\left(\bar{x}\right)_i^*\left(\bar{x}\right)_k\right]\right\}\\
&= \sum_{k=1}^{M}\sum_{l=1}^{M}E\left[\left(\bar{x}\right)_i^*\left(\bar{x}\right)_k\right]E\left[\left(\bar{y}\bar{y}^H\right)_{kl}\right]E\left[\left(\bar{x}\right)_l^*\left(\bar{x}\right)_j\right]\\
&\quad+\sum_{k=1}^{M}\sum_{l=1}^{M}E\left[\left(\bar{x}\right)_i^*\left(\bar{x}\right)_j\right]E\left[\left(\bar{y}\bar{y}^H\right)_{kl}\right]E\left[\left(\bar{x}\right)_l^*\left(\bar{x}\right)_k\right]\\
&= \sum_{k=1}^{M}\sum_{l=1}^{M}\left(\mathbf{R}_x\right)_{ik}\left(\mathbf{R}_y\right)_{kl}\left(\mathbf{R}_x\right)_{lj} + \left(\mathbf{R}_x\right)_{ij}\mathrm{Tr}\left(\mathbf{R}_x\mathbf{R}_y\right)\\
&= \left(\mathbf{R}_x\mathbf{R}_y\mathbf{R}_x\right)_{ij} + \left(\mathbf{R}_x\mathrm{Tr}\left(\mathbf{R}_x\mathbf{R}_y\right)\right)_{ij}.
\end{aligned}$$

Applying the above identity, the required quantities are

$$\begin{aligned}
&E\left[\bar{x}_k^*\bar{x}_k^T\tilde{\Delta}_{k-1}^{\boldsymbol{l}}\left(\tilde{\Delta}_{k-1}^{\boldsymbol{l}}\right)^H\bar{x}_k^*\bar{x}_k^T\right]\\
&= \mathbf{\Phi}_k^x\mathbf{Z}_k\mathbf{\Phi}_k^x + \mathbf{\Phi}_k^x\mathrm{Tr}\left[\mathbf{\Phi}_k^x\mathbf{Z}_k\right]
\end{aligned}$$

and

$$\begin{aligned}
&E\left\{\bar{x}_k^*\bar{x}_k^T E\left[\bar{\Delta}_{k-1}^{\boldsymbol{l}}\right]E\left[\bar{\Delta}_{k-1}^{\boldsymbol{l}}\right]^H\bar{x}_k^*\bar{x}_k^T\right\}\\
&= \mathbf{\Phi}_k^x E\left[\bar{\Delta}_{k-1}\right]E\left[\bar{\Delta}_{k-1}\right]^H\mathbf{\Phi}_k^x\\
&\quad+E\left[\bar{\Delta}_{k-1}\right]^H\mathbf{\Phi}_k^x E\left[\bar{\Delta}_{k-1}\right]\mathbf{\Phi}_k^x.
\end{aligned}$$

REFERENCES

[1] O. Macchi, *Adaptive Processing: The Least Mean Square Approach With Applications in Transmission*. New York: Wiley, 1995.

[2] S. Haykin, *Adaptive Filter Theory*, 3rd ed. Englewood Cliffs, NJ: Prentice-Hall, 1996.

[3] O. Macchi and N. J. Bershad, "Adaptive recovery of a chirped sinusoidal in noise: I. Performance of the RLS algorithm," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 39, pp. 583–594, Mar. 1991.

[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] N. J. Bershad and O. Macchi, "Adaptive recovery of a chirped sinusoidal in noise: II. Performance of the LMS algorithm," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 39, pp. 595–602, Mar. 1991.

[6] O. Macchi, N. J. Bershad, and M. Mboup, "Steady-state superiority of LMS over LS for time-varying line enhancer in noisy environment," in *Proc. Inst. Elect. Eng. F*, vol. 138, Aug. 1991, pp. 354–360.

[7] S. Haykin, A. Sayed, J. Zeidler, P. Yee, and P. Wei, "Adaptive tracking of linear time-variant systems by extended RLS algorithms," *IEEE Trans. Signal Processing*, vol. 45, pp. 1118–1127, May 1997.

[8] P. Wei, J. R. Zeidler, and W. H. Ku, "Adaptive recovery of a chirped signal using the RLS algorithm," *IEEE Trans. Signal Processing*, vol. 45, pp. 363–376, Feb. 1997.

[9] P. C. Wei, "Performance evaluation of adaptive filtering algorithms for the mobile communication environment," Ph.D. dissertation, Univ. California at San Diego, La Jolla, 1995.

Paul C. Wei (M'95) received the B.S.E.E. degree from the University of California, Los Angeles, in 1989 and the M.S.E.E. and Ph.D.E.E. degrees from the University of California, San Diego (UCSD), La Jolla, in 1991 and 1995, respectively.

He is currently a Staff Systems Engineer with Via Telecom, San Diego, where he has worked on designs for GSM, IS95, and IS2000 systems. He was an intern with the GM Technical Center, Warren, MI, during the summers of 1988 to 1990. From 1991 to 1995, he was a Research Assistant with the Industry/University Cooperative Research Center on Ultra High Speed Integrated Circuits and Systems, UCSD. His research interests include applications of signal processing techniques to broadband communications and performance of adaptive algorithms in high-speed hardware implementations.

Jun Han (S'98) received the B.Eng. degree in applied electronic technique from the University of Petroleum of East China and the M.Eng. degree in data transmission and processing from the University of Petroleum, Beijing, China, in 1989 and 1992, respectively. He is currently pursuing the Ph.D. degree with the Department of Electrical and Computer Engineering, University of California, San Diego, La Jolla.

His research interests include digital signal processing in wireless communications and performance analysis of adaptive filters.

James R. Zeidler (M'76–SM'84–F'94) has been a Scientist at the Space and Naval Warfare Systems Center, San Diego, CA, since 1974. He has also been an Adjunct Professor with the Electrical and Computer Engineering Department, University of California, San Diego, La Jolla, since 1988.

His current research interests are in adaptive signal processing, communications signal processing, and wireless communication networks.

Dr. Zeidler was an Associate Editor of the IEEE TRANSACTIONS ON SIGNAL PROCESSING from 1991 to 1994. He was co-recipient of the award for best unclassified paper at the IEEE Military Communications Conference in 1995 and received the Lauritsen–Bennet Award for achievement in science in 2000 and the Navy Meritorious Civilian Service Award in 1991.

Walter H. Ku (M'91) received the B.S. degree (with honors) in electrical engineering from the Moore School of Electrical Engineering, University of Pennsylvania, Philadelphia, and the M.S. and Ph.D. degrees in electrical engineering from the Polytechnic Institute of Brooklyn, Brooklyn, NY.

In 1977, he was the first occupant of the Naval Electronic Systems Command (NAVELEX) Research Chair Professorship at the Naval Post-graduate school, Monterey, CA. As the NAVELEX Research Chair holder, he was an Expert Consultant to NAVELEX (now SPAWAR), Naval Research Laboratory (NRL), and OUSDRE. He has served as a Consultant to the Department of Defense (DDR&E and ARPA) on the VHSIC and various GaAs monolithic integrated circuits programs, including the first DARPA-funded C- and X-band space-based radar modules, Air Force RADC, Griffith AFB, Rome, NY, and industrial laboratories. Since September 1985, he has been Professor of Electrical and Computer Engineering at the University of California, San Diego (UCSD), La Jolla, and is the Founding Director of the NSF I/UC Research Center on Ultra-High Speed Integrated Circuits and Systems (ICAS). He is also the Principal Investigator of a new five-year grant from DDR&E Focused Research Initiative (FRI) on Broadband Wireless Multimedia (BWM) Communications Systems. This FRI program was awarded in January 1995 and is a consortium led by UCSD with Advanced Digital Technology (ADTR), Hughes Raythean Rockwell, and TRW as team members.

Dr. Ku is a member of Eta Kappa Nu and Tau Beta Pi.