

Nonlinear Effects of the LMS Adaptive Predictor for Chirped Input Signals

Jun Han

*Department of Electrical and Computer Engineering, University of California, San Diego, La Jolla, CA 92093-0407, USA
Email: junhan@ece.ucsd.edu*

James R. Zeidler

*Department of Electrical and Computer Engineering, University of California, San Diego, La Jolla, CA 92093-0407, USA
Email: zeidler@ece.ucsd.edu*

*Space and Naval Warfare Center, D8505, San Diego, CA 92152, USA
Email: zeidler@spawar.navy.mil*

Walter H. Ku

*Department of Electrical and Computer Engineering, University of California, San Diego, La Jolla, CA 92093-0407, USA
Email: ku@ece.ucsd.edu*

Received 30 July 2001 and in revised form 11 October 2001

This paper investigates the nonlinear effects of the Least Mean Square (LMS) adaptive predictor. Traditional analysis of the adaptive filter ignores the statistical dependence among successive tap-input vectors and bounds the performance of the adaptive filter by that of the finite-length Wiener filter. It is shown that the nonlinear effects make it possible for an adaptive transversal prediction filter to significantly outperform the finite-length Wiener predictor. An approach is derived to approximate the total steady-state Mean Square Error (MSE) for LMS adaptive predictors with stationary or chirped input signals. This approach shows that, while the nonlinear effect is small for the one-step LMS adaptive predictor, it increases in magnitude as the prediction distance is increased. We also show that the nonlinear effect of the LMS adaptive predictor is more significant than that of the Recursive Least Square adaptive predictor.

Keywords and phrases: adaptive filter, linear prediction, least mean square, recursive least square, tracking, autoregressive model.

1. INTRODUCTION

The Least Mean Square (LMS) adaptive filter is widely used in many applications partly due to the simplicity of its implementation [1]. The simplicity belies the fact that the adaptive LMS filter is a complex nonlinear estimator [2, 3, 4, 5, 6, 7, 8, 9]. Traditional analysis of adaptive filter performance is restricted to a statistical analysis of the LMS algorithm under a set of independence assumptions that ignore the statistical dependence among successive tap-input vectors [1]. The Mean Square Error (MSE) of the LMS adaptive filter using these assumptions is bounded by that of the corresponding finite-length Wiener filter, and the MSE of the adaptive filter increases monotonically as a function of the adaptation step-size. While simulations show that this simplified analysis predicts the performance reasonably well in many applications for small step-size, it was shown that in some applications there is a large discrepancy between the

simulation results and what the independence analysis predicts [2, 3, 4, 5, 6, 7, 8, 9]. The reason for the discrepancy is that these well-known assumptions mask the nonlinear effects that arise in LMS adaptive filters. It has been shown that it is possible for the LMS adaptive filter to outperform the finite-length Wiener filter in MSE for the cases of adaptive channel equalization for sinusoidal and first-order autoregressive process (AR1) interference suppression [9], and adaptive noise cancellation for narrowband AR1 signals when the primary and reference signals have slightly different frequencies [7]. An error transfer function approach is also derived in [9] to give an approximate expression for the total steady-state MSE of the LMS adaptive channel equalizer.

In this paper, the nonlinear effects in a third application of adaptive filters, adaptive prediction, is studied. The class of input signals which will be considered for adaptive prediction are the stationary and chirped narrowband input signals for varying chirp rates and bandwidth. This class of signals has

been used to represent a signal whose spectrum is frequency offset and shifted with time in a nonstationary mobile communications environment [10, 11]. They are different from those considered in [2, 3, 4, 5, 6, 7, 8, 9] because they have a time-varying Power Spectral Density (PSD). Since they do not have a fixed PSD, the error transfer function approach [9] is not directly applicable. However, since the chirped signal has a constant spectral shifting rate, this special class of nonstationary inputs can be analyzed as stationary inputs by an unchirped transform defined below. It is proven in this paper that the MSE of the standard LMS adaptive predictor with a chirped input signal is equal to the MSE of a transformed LMS adaptive predictor with the corresponding stationary input signal. An error transfer function approach is derived for the transformed LMS algorithm with stationary input signals so as to approximate the MSE of chirped signal prediction. To bound the performance of the LMS adaptive predictor, the MSE of the optimal estimator (the infinite-length one-step causal Wiener predictor) is calculated.

To compare the magnitude of nonlinear effects of the LMS and RLS adaptive predictors, the error feedback transfer function is also derived for the RLS algorithm. By comparing the contributions of past errors to the current estimates in the two algorithms, it is shown that the LMS algorithm uses information from past prediction errors more effectively than the RLS algorithm.

2. BACKGROUND

The adaptive predictor application considered is the adaptive recovery of narrowband signals from embedded Additive White Gaussian Noise (AWGN). The narrowband input signal is modeled as an AR1 process. It is shown in [11] that the AR1 process provides a reasonable approximation to a BPSK communication signal. The AR1 process satisfies the recursive equation

$$s_n = a s_{n-1} + v_n, \quad (1)$$

where v_n is a white noise process, with $\sigma_v^2 = P_s(1 - |a|^2)$ and P_s is the power of the AR1 process. The corresponding chirped AR1 signal s_n^c , where superscript c denotes the chirped signal, has the following form [11]:

$$s_n^c = a\Omega\Psi^{n-(1/2)}s_{n-1}^c + v_n^c, \quad (2)$$

where $\Omega = e^{j\omega_0}$, ω_0 defines the initial center frequency of the spectrum, $\Psi = e^{j\psi}$, ψ is the chirp rate which linearly shifts the center frequency with time, and v_n^c is a white noise process with the same statistics as v_n . This chirped AR1 signal can be used to represent a signal whose spectrum is frequency offset and shifted with time in a nonstationary mobile communications environment. The chirped AR1 process and chirped sinusoid have been used to study the tracking behavior of adaptive filters because they provide an input signal with a single constant nonstationary component [11, 12, 13, 14, 15]. Chirped signals are also used in conjunction with OFDM communications and radar systems to optimize power transmission over a wide bandwidth when

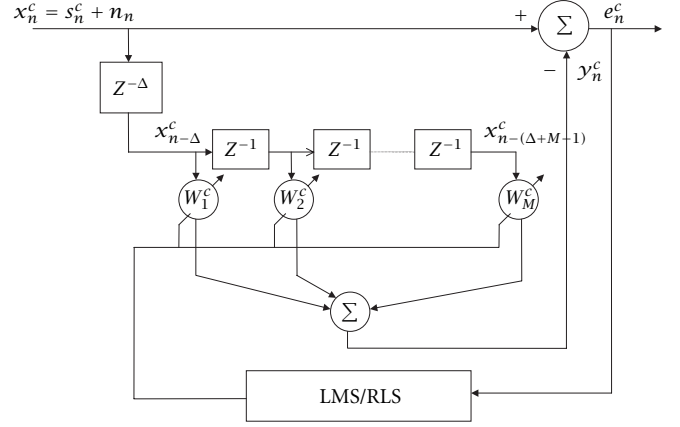


FIGURE 1: LMS/RLS Δ -step predictor structure.

the propagation medium is time-varying [10].

At the receiver the signal is given by

$$x_n^c = s_n^c + n_n, \quad (3)$$

where n_n is the AWGN process with power P_n .

Figure 1 represents the linear Δ -step adaptive predictor structure to be analyzed, where $W^c(n)$ are the adaptive filter weights. The weight update equation of the LMS algorithm is

$$W^c(n+1) = W^c(n) + \mu X^{c*}(n)e_n^c, \quad (4)$$

where μ is the step-size parameter of the adaptive algorithm, $X^c(n)$ is the adaptive filter input tap-vector at time n , and $*$ denotes the complex conjugate. For the Δ -step predictor,

$$X^c(n) = \begin{bmatrix} x_{n-\Delta}^c \\ x_{n-(\Delta+1)}^c \\ \vdots \\ x_{n-(\Delta+M-1)}^c \end{bmatrix}. \quad (5)$$

The error update equation is given by

$$e_{n+1}^c = x_{n+1}^c - W^{cT}(n+1)X^c(n+1). \quad (6)$$

The finite-length Wiener predictor weight and the corresponding MSE are given as

$$\begin{aligned} W_0^c(n) &= [\mathbf{R}^c(n)]^{-1}P^c(n), \\ J_w^c(n) &= P_s + P_n - P^c(n)^H W_0^c(n), \end{aligned} \quad (7)$$

where $\mathbf{R}^c(n)$ is the autocorrelation matrix of the input signal vector $\mathbf{R}^c(n) = E[X^{c*}(n)X^c(n)]$, $P^c(n)$ is the cross-correlation of the input signal vector with the desired response $P^c(n) = E[X^{c*}(n)x_n^c]$, and $J_w^c(n) = E[|x_n^c - [W_0^c(n)]^T X^c(n)|^2]$ is the MSE of the finite-length Wiener predictor. By setting $\omega = 0$ and $\psi = 0$, \mathbf{R} is the autocorrelation matrix of the corresponding stationary baseband input signal x_n , P is the cross-correlation vector, Wiener predictor is $W_0 = \mathbf{R}^{-1}P$, and the Wiener MSE is J_w . It has been shown

[11] that $J_w^c(n)$ of the Wiener predictor for the chirped input signal x_n^c is equal to J_w of the Wiener predictor for the corresponding stationary baseband input signal x_n .

3. THE LMS PREDICTOR FOR CHIRPED INPUT SIGNALS

The error transfer function approach derived in [9] provides a method to approximate the total steady-state MSE of the LMS adaptive filter without explicitly invoking the independence assumptions for wide-sense stationary input signals, that is, signals with a fixed PSD. For a chirped input signal x_n^c , the PSD is constantly shifting with time, and this approach is not directly applicable. However, the adaptive recovery of a narrowband chirped signal using a Δ -step transversal predictor has one important characteristic, that is, the frequency offset among the input signal taps is the chirp rate ψ , and the frequency offset between the desired response x_n^c and input signal vector $X^c(n)$ is $\Delta \times \psi$. By multiplying the chirped input signal by a negative frequency offset sequence, we can transform the chirped signal s_n^c to its stationary form s_n and leave the noise component n_n unchanged since the AWGN has a constant spectral envelope across all frequencies. In the following, it is shown that the above transform will not change the MSE of the LMS adaptive predictor for a chirped input signal. This allows the error transfer function approach to be applied to rotated LMS algorithm with the transformed input signals in order to approximate the MSE of the standard LMS adaptive predictor with chirped input signals.

3.1. Equivalence of MSEs

For a chirped input signal $x_n^c = s_n^c + n_n$, $n = 0, 1, 2, \dots$, where s_n^c has initial center frequency ω_0 and chirp rate ψ , we define a transformed process,

$$x_n^u = \Omega^{-n} \Psi^{-n^2/2} x_n^c, \quad n = 0, 1, 2, \dots, \quad (8)$$

where superscript u denotes the unchirped process. This operation will transform the chirped input signal to a stationary baseband signal, and it will change the formulation of the standard LMS algorithm in (4) and (6).

Multiplying (6) by $\Omega^{-(n+1)} \Psi^{-(n+1)^2/2}$, and defining $e_n^u = \Omega^{-n} \Psi^{-n^2/2} e_n^c$, $n = 0, 1, 2, \dots$ (which is the transformed version of the predictor error signal for chirped input process using the LMS adaptive predictor), this transforms (6) to

$$e_{n+1}^u = x_{n+1}^u - W^{uT}(n+1)X^u(n+1). \quad (9)$$

In (9),

$$\begin{aligned} W^u(n+1) &= \text{diag} \left\{ \Omega^{-\Delta} \Psi^{-(n+1)^2/2 + [n-(\Delta-1)]^2/2}, \right. \\ &\quad \Omega^{-(\Delta+1)} \Psi^{-(n+1)^2/2 + [n-\Delta]^2/2}, \\ &\quad \vdots \\ &\quad \left. \Omega^{-(\Delta+M-1)} \Psi^{-(n+1)^2/2 + [n-(\Delta+M-2)]^2/2} \right\} W^c(n+1) \end{aligned} \quad (10)$$

is the corresponding predictor weight in the transformed domain and

$$X^u(n+1) = \begin{bmatrix} \Omega^{-[n-(\Delta-1)]} \Psi^{-[n-(\Delta-1)]^2/2} x_{n-(\Delta-1)}^c \\ \Omega^{-(n-\Delta)} \Psi^{-(n-\Delta)^2/2} x_{n-\Delta}^c \\ \vdots \\ \Omega^{-[n-(\Delta+M-2)]} \Psi^{-[n-(\Delta+M-2)]^2/2} x_{n-(\Delta+M-2)}^c \end{bmatrix} \quad (11)$$

is the transformed version of the chirped input signal vector $X^c(n+1)$. Applying (8) to the vector elements in $X^c(n+1)$ results in stationary baseband signals.

Using (10) and (11), (4) can be shown to become

$$W^u(n+1) = \mathbf{V}_\Delta^* [W^u(n) + \mu X^u(n) e_n^u], \quad (12)$$

where

$$\mathbf{V}_\Delta \triangleq \Psi^{\Delta-1} \text{diag} \{ \Psi^1, \Psi^2, \dots, \Psi^M \} \quad (13)$$

is the chirp rotation matrix. Since e_n^u is the transformed version of e_n^c , they have the same power, that is,

$$E[|e_n^u|^2] = E[|\Omega^{-n} \Psi^{-n^2/2} e_n^c|^2] = E[|e_n^c|^2]. \quad (14)$$

Consequently, the MSE of the LMS adaptive predictor with a chirped input signal x_n^c is equal to the MSE of a different LMS adaptive predictor with a corresponding stationary baseband input signal x_n^u . Note that the two adaptive predictors have the same length M and step-size μ . Equations (9) and (12) define the error and weight vectors of the rotated LMS adaptive predictor. The only difference between these equations and the standard LMS adaptive predictor for stationary input signals as in (4) and (6) is that the weight vector is rotated in frequency by the chirp matrix \mathbf{V}_Δ after each normal LMS update, as shown in Figures 2 and 3.

3.2. Error transfer function approach for the rotated LMS adaptive predictor

First, we decompose the rotated LMS adaptive predictor weight into the sum of a time-invariant finite-length Wiener

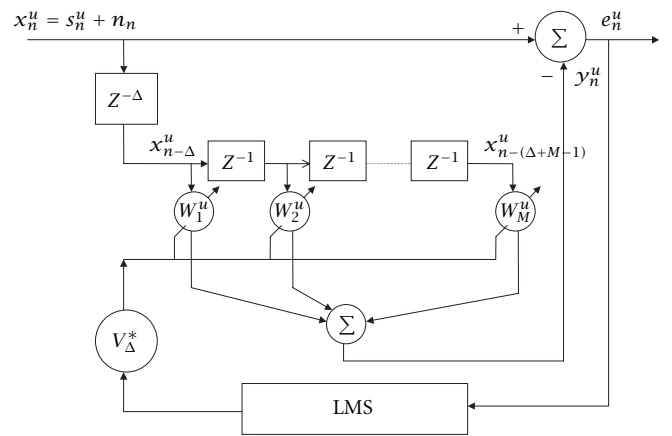


FIGURE 2: Rotated LMS Δ -step predictor structure.

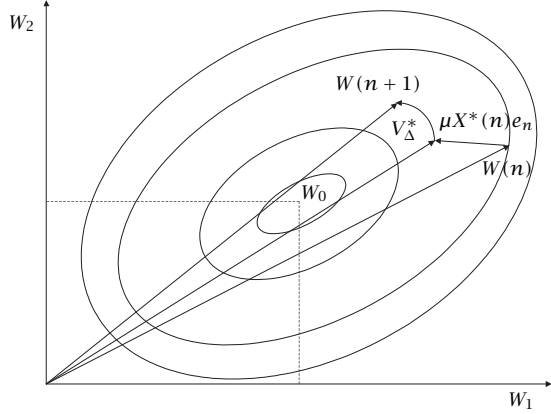


FIGURE 3: Rotation of weight updates in rotated LMS adaptive predictor.

predictor weight and a time-varying misadjustment component

$$W^u(n) = W_0 + W_{\text{mis}}^u(n). \quad (15)$$

$W_{\text{mis}}^u(n)$ is further decomposed as

$$W_{\text{mis}}^u(n) = \bar{W}_{\text{mis}}^u(n) + \tilde{W}_{\text{mis}}^u(n), \quad (16)$$

where $\bar{W}_{\text{mis}}^u(n) = E[W_{\text{mis}}^u(n)]$ is the mean weight misadjustment corresponding to the weight fluctuation caused by weight rotation. From (12), the weight misadjustment is given by

$$W_{\text{mis}}^u(n+1) = \mathbf{V}_{\Delta}^* \left\{ [\mathbf{I} - \mu X^{u*}(n) X^{uT}(n)] W_{\text{mis}}^u(n) + \mu X^{u*}(n) e_0 \right\} - (\mathbf{I} - \mathbf{V}_{\Delta}^*) W_0, \quad (17)$$

where \mathbf{I} is the identity matrix. The mean weight misadjustment is

$$\bar{W}_{\text{mis}}^u(n+1) = \mathbf{V}_{\Delta}^* (\mathbf{I} - \mu \mathbf{R}) \bar{W}_{\text{mis}}^u(n) - (\mathbf{I} - \mathbf{V}_{\Delta}^*) W_0, \quad (18)$$

when $n \rightarrow \infty$, that is, the adaptive filter reaches steady state,

$$\bar{W}_{\text{mis}}^u = \bar{W}_{\text{mis}}^u(\infty) = -(\Lambda + \mu \mathbf{R})^{-1} \Lambda W_0, \quad (19)$$

where $\Lambda \triangleq \mathbf{V}_{\Delta} - \mathbf{I}$. Note that in (18), it is assumed that $\bar{W}_{\text{mis}}^u(n)$ is independent of $X^{u*}(n) X^{uT}(n)$, and it is not necessary for $\bar{W}_{\text{mis}}^u(n)$ to be independent of $X^u(n)$. This steady-state mean weight misadjustment term corresponds to the lag weight misadjustment of the LMS adaptive predictor with a chirped input process as shown in [12].

The recursive weight update equation (12) can be written as

$$\begin{aligned} W^u(n) &= \mathbf{V}_{\Delta}^* [W^u(n-1) + \mu X^{u*}(n-1) e_{n-1}^u] \\ &= \mathbf{V}_{\Delta}^{*n} W^u(0) + \mu \sum_{j=0}^{n-1} \mathbf{V}_{\Delta}^{*(n-j)} X^{u*}(j) e_j^u. \end{aligned} \quad (20)$$

The adaptive filter output is

$$y_n^u = W^{uT}(0) \mathbf{V}_{\Delta}^{*n} X^u(n) + \mu \sum_{j=0}^{n-1} e_j^u X^{uH}(j) \mathbf{V}_{\Delta}^{*(n-j)} X^u(n). \quad (21)$$

At steady state, $\mathbf{V}_{\Delta}^{*n} W^u(0)$ can be replaced with $W_0 + \bar{W}_{\text{mis}}^u$, thus the error process e_n^u satisfies the recursive difference equation

$$\begin{aligned} e_n^u + \mu \sum_{j=0}^{n-1} e_j^u X^{uH}(j) \mathbf{V}_{\Delta}^{*(n-j)} X^u(n) \\ = x_n^u - [W_0 + \bar{W}_{\text{mis}}^u]^T X^u(n). \end{aligned} \quad (22)$$

Using the approximations [9]

$$\begin{aligned} X^{uH}(j) X^u(n) &\approx M r_x^u(n-j), \\ \mathbf{V}_{\Delta} &\approx \Psi^{\Delta+(M-1)/2} \mathbf{I}, \quad \psi \ll 1, \end{aligned} \quad (23)$$

where $r_x^u(k)$ is the autocorrelation of the stationary input signal x_n^u , we have

$$X^{uH}(j) \mathbf{V}_{\Delta}^{*(n-j)} X^u(n) \approx M r_x^c(n-j), \quad (24)$$

where

$$r_x^c(n-j) \triangleq \Psi^{-(\Delta+(M-1)/2)(n-j)} r_x^u(n-j). \quad (25)$$

Equation (22) can be approximated by a standard difference equation with constant coefficients as

$$e_n^u + \mu M \sum_{j=0}^{n-1} r_x^c(n-j) e_j^u = x_n^u - [W_0 + \bar{W}_{\text{mis}}^u]^T X^u(n). \quad (26)$$

The left-hand side of (26) is the convolution of $[e_n^u, e_{n-1}^u, e_{n-2}^u, \dots, e_0^u]$ with $[1, \mu M r_x^c(1), \mu M r_x^c(2), \dots, \mu M r_x^c(n)]$. We can interpret the steady-state ($n \rightarrow \infty$) rotated LMS adaptive predictor error e_n^u as the output of a time-invariant linear system with transfer function $H(z)$ driven by the wide-sense stationary error process $x_n^u - [W_0 + \bar{W}_{\text{mis}}^u]^T X^u(n)$, where $H(z)$ is given by

$$H(z) = \frac{1}{1 + \mu M R(z)}, \quad (27)$$

$$R(z) = \sum_{m=1}^{\infty} r_x^c(m) z^{-m}. \quad (28)$$

The steady-state MSE of the rotated LMS adaptive predictor is thus

$$J_{\text{lms}} = \frac{1}{2\pi j} \oint_{|z|=1} |H(z)|^2 |1 - W_0(z) - \bar{W}_{\text{mis}}^u(z)|^2 S_{xx}^u(z) \frac{dz}{z}, \quad (29)$$

where

$$\begin{aligned} W_0(z) &= \sum_{j=\Delta}^{M+\Delta-1} W_0(j) z^{-j}, \\ \bar{W}_{\text{mis}}^u(z) &= \sum_{j=\Delta}^{M+\Delta-1} \bar{W}_{\text{mis}}^u(j) z^{-j} \end{aligned} \quad (30)$$

are the transfer functions of the finite-length Wiener predictor and mean weight misadjustment of the rotated LMS adaptive predictor, respectively. $S_{x_x}^u(z)$ is the PSD of the stationary input process x_n^u transformed from the chirped input signal x_n^c .

The error transfer function approach can also be applied to the Normalized LMS (NLMS) algorithm as defined below [9]

$$W^c(n+1) = W^c(n) + \frac{\mu}{\|X^c(n)\|^2} X^{c*}(n) e_n^c \quad (31)$$

with

$$H(z) = \frac{1}{1 + \mu R(z)/(P_s + P_n)}. \quad (32)$$

4. BOUND OF THE Δ -STEP ADAPTIVE PREDICTOR

Using the recursive equations (4) and (6), it follows from [8] that the LMS adaptive predictor is a nonlinear estimator of x_n^c . The estimate is a function of all the past samples it used in the recursion: $\{x_{n-1}^c, x_{n-2}^c, \dots, x_{-\infty}^c\}$. Denoting C_{lms} as the LMS estimator, we have

$$\hat{x}_n = C_{\text{lms}}[x_{n-1}^c, x_{n-2}^c, \dots, x_{-\infty}^c] \quad (33)$$

and the estimation MSE is given by $J_{\text{lms}} = E[|e_n^c|^2]$.

The optimal MSE estimator C_{opt} using the same data as the adaptive predictor is given by

$$\begin{aligned} y_n &= C_{\text{opt}}[x_{n-1}^c, x_{n-2}^c, \dots, x_{-\infty}^c] \\ &= E[x_n^c | x_{n-1}^c, x_{n-2}^c, \dots, x_{-\infty}^c]. \end{aligned} \quad (34)$$

For wide-sense stationary input process, the performance of Δ -step prediction is bounded by that of the optimal MSE estimator, which is the one-step infinite-length Wiener predictor. The optimal estimator is independent of the prediction distance Δ .

Since the finite-length Wiener predictor is not recursive, it can be written as

$$\hat{x}_n = E[x_n^c | x_{n-\Delta}^c, x_{n-(\Delta+1)}^c, \dots, x_{n-(\Delta+M-1)}^c]. \quad (35)$$

To illustrate that the nonlinear effect is small for the one-step LMS adaptive predictor, but increases in magnitude as the prediction distance Δ is increased, Figures 4 and 5 delineate the data utilized by the adaptive predictor, the finite-length Wiener predictor and the optimal estimator for one-step and Δ -step prediction ($\Delta > 1$). Figure 4 shows that for one-step prediction of x_n^c , the data available to the adaptive predictor and the optimal estimator but not available to the finite-length Wiener predictor is defined by the sequence $[x_{n-M}^c, x_{n-(M+1)}^c, \dots, x_{-\infty}^c]$. The contribution of this signal segment to the prediction of x_n^c is negligible because the correlation of the desired signal and the data segment is small. In contrast, for multiple-step prediction shown in Figure 5, the additional data which is available to the adaptive predictor and the optimal es-

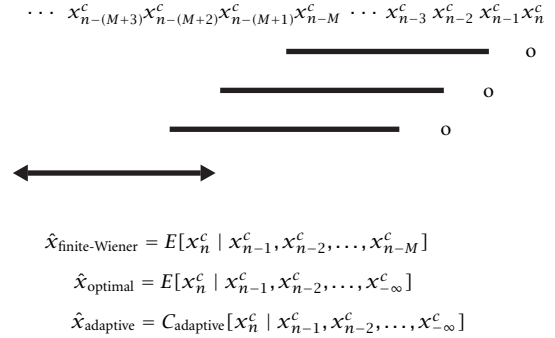


FIGURE 4: Information utilized by one-step adaptive predictor, finite-length Wiener predictor, and optimal estimator. The data segment marked by arrows is the information available to adaptive predictor and optimal estimator, but not available to finite-length Wiener predictor in the prediction of x_n^c .

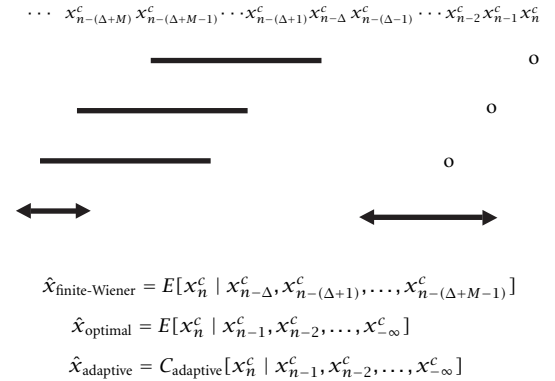


FIGURE 5: Information utilized by Δ -step adaptive predictor, finite-length Wiener predictor, and optimal estimator. The data segments marked by arrows are the information available to adaptive predictor and optimal estimator, but not available to the finite-length Wiener predictor in the prediction of x_n^c .

imator but not available to the finite-length Wiener predictor has two components, $[x_{n-1}^c, x_{n-2}^c, \dots, x_{n-(\Delta-1)}^c]$ and $[x_{n-(\Delta+M)}^c, x_{n-(\Delta+M+1)}^c, \dots, x_{-\infty}^c]$. The main contribution to the prediction of x_n^c is the first term, because for a narrow-band signal, the correlation of the first term with x_n^c is much larger than the correlation of x_n^c with the second term. Note also that the correlation of x_n^c with the first component is also larger than the correlation of x_n^c with the predictor input signal $[x_{n-\Delta}^c, x_{n-(\Delta+1)}^c, \dots, x_{n-(\Delta+M-1)}^c]$. With an increase in the prediction distance Δ , there will be more information available to the adaptive predictor than to the finite-length Wiener predictor, and consequently the adaptive predictor may outperform the finite-length Wiener predictor. Conversely, the adaptive predictor performance is bounded by the one-step infinite-length Wiener predictor since they utilize the same amount of information and there is misadjustment noise associated with the adaptive predictor. Note that Figures 4 and 5 are also applicable to RLS adaptive predictors, that is, for one-step and multiple-step predictions. The LMS and RLS adaptive predictors use information from the same input data, so that any performance difference between these

two algorithms must be explained from their difference in adjusting the filter weights according to the feedback errors.

5. THE COMPARATIVE PERFORMANCE OF THE LMS AND RLS ALGORITHMS

For simplicity, we only compare the two adaptive algorithms with stationary input signals x_n . The weight update equation of the exponentially weighted RLS adaptive algorithm is given by [1]

$$W(n+1) = W(n) + \Phi^{-1}(n)X^*(n)e_n, \quad (36)$$

where $\Phi(n) = \sum_{i=0}^n \lambda^{n-i} X^*(i)X(i)^T$ is the input signal autocorrelation matrix estimate at time n , and λ is the forgetting factor of the RLS algorithm. Decompose the weight vector as

$$W(n) = W_0 + W_{\text{mis}}(n). \quad (37)$$

The predictor error is

$$\begin{aligned} e_n &= x_n - W^T(n)X(n) \\ &= x_n - W_0^T X(n) - \sum_{j=0}^{n-1} e_j X^H(j)\Phi(j)X(n). \end{aligned} \quad (38)$$

The steady-state error update equation of the RLS adaptive predictor is given by

$$e_n + \sum_{j=0}^{n-1} e_j X^H(j)\Phi^{-1}(j)X(n) = x_n - W_0^T X(n). \quad (39)$$

The following two approximations are used at steady state:

$$\begin{aligned} \Phi^{-1}(j) &\approx (1-\lambda)\mathbf{R}^{-1}, \\ X^H(j)\Phi^{-1}(j)X(n) &\approx (1-\lambda)\text{trace}(\mathbf{R}^{-1}E[X(n)X^H(j)]). \end{aligned} \quad (40)$$

Defining $\text{cr}(j) = (1-\lambda)\text{trace}(\mathbf{R}^{-1}E[X(n)X^H(n-j)])$, (39) becomes

$$e_n + \sum_{j=1}^{\infty} e_{n-j} \text{cr}_j = x_n - W_0^T X(n). \quad (41)$$

To show the difference of the two algorithms in utilizing past prediction errors, the error feedback equation of the LMS algorithm (26) is rewritten for a stationary input signal as

$$e_n + \sum_{j=1}^{\infty} e_{n-j} \text{cl}_j = x_n - W_0^T X(n), \quad (42)$$

where $\text{cl}_j = \mu M r_x(j)$. Figure 6 is a plot of cl_j and cr_j , $j = 1, 2, \dots, 50$ for a narrowband AR1 input signal embedded in AWGN, with AR1 pole location $a = 0.99$, SNR = 10 dB, adaptive filter length $M = 25$. The LMS step-size $\mu = 0.01$, and the RLS forgetting factor $\lambda = 0.9$. The error feedback coefficients of the RLS adaptive predictor exhibit a null for small j , which means that the contributions from the most

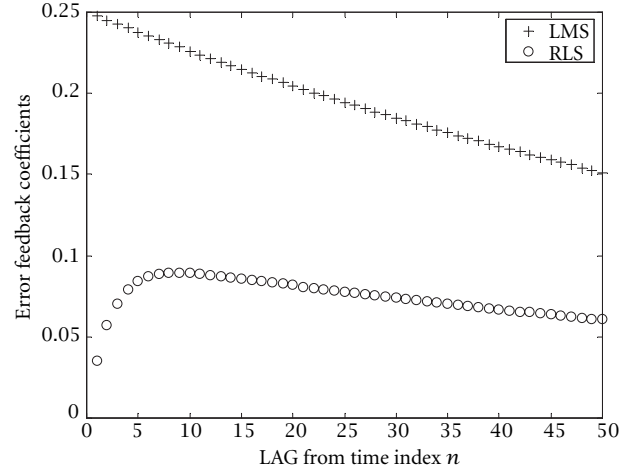


FIGURE 6: Error feedback coefficients of the LMS and RLS adaptive predictors ($\mu = 0.01, \lambda = 0.9$).

recent prediction errors to the current estimate at time index n are nulled out. For the LMS adaptive predictor the most recent prediction errors contribute more to the current estimate than the time delayed prediction errors. Note that the choices of μ and λ only affect the magnitudes, not the shapes of the curves.

6. SIMULATIONS

For a chirped AR1 input, the autocorrelation of input signal vectors $r_x^c(k)$ is given by

$$r_x^c(j) = \Psi^{-(\Delta+(M-1)/2)j} P_s a^j, \quad (43)$$

where a is pole location of the transformed stationary baseband AR1 input signal. Equation (28) becomes

$$R(z) = \frac{P_s a_c z^{-1}}{1 - a_c z^{-1}}. \quad (44)$$

The feedback transfer function in (27) is thus

$$H(z) = \frac{1 - a_c z^{-1}}{1 - g_c z^{-1}}, \quad (45)$$

where $a_c = a\Psi^{-(\Delta+(M-1)/2)}$, $g_c = (1 - \mu M P_s) a_c$. The steady-state MSE approximation for the LMS adaptive predictor can be computed from (29) using (45). Similarly, the steady-state MSE approximation of the NLMS adaptive predictor can be calculated from (29), (32), and (43). By setting $\psi = 0$, the MSE approximation of the standard LMS adaptive predictor for a stationary input signal is calculated.

In the following simulations for multiple-step prediction, the NLMS adaptive predictors are used instead of the standard LMS adaptive predictors because the NLMS algorithm is stable for relatively larger values of the adaptive filter step-size ($0 < \mu < 2$) [9], where the nonlinear effects of adaptive algorithm are most significant. The MSEs of the finite-length Wiener predictor and the optimal estimator are calculated theoretically for AR1 input processes.

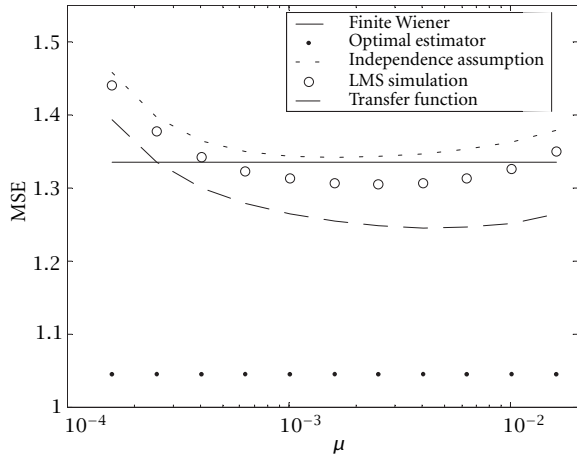


FIGURE 7: Comparison of MSEs of one-step LMS adaptive predictor with very narrowband input signal as a function of adaptation constant μ . $a = 0.999$, $M = 2$, $\text{SNR} = 1$, chirp rate $\psi = 5\pi e - 5$.

Figure 7 is a plot of MSEs for one-step LMS adaptive predictors as a function of filter step-size μ with a chirped AR1 input signal, where signal initial frequency $\omega_0 = 0.2\pi$, chirp rate $\psi = 5\pi e - 5$, AR1 process pole location $\alpha = 0.999$, signal power $P_s = 1$, $\text{SNR} = 0$ dB, filter length $M = 2$. Simulation results and theoretical calculations using both the transfer function approach and the independence assumptions are plotted. These results are compared to the MSEs obtained for the finite-length Wiener predictor and the optimal estimator. It can be seen that in a small range of adaptive filter step-size parameters μ , the MSE from the error transfer function approach and simulation results are smaller than the MSE of the finite-length Wiener predictor. Extensive simulations and analytical results show that for the one-step LMS adaptive predictor, the nonlinear effect is small and observable only for very small filter length, very narrow bandwidth input signals. One possible explanation of this phenomenon is that under these conditions, the information in $\{x_{n-(M+1)}^c, x_{n-(M+2)}^c, \dots, x_\infty^c\}$ which is available to adaptive predictor but not available to finite-length Wiener predictor, will have effective contributions to the prediction of current signal x_n^c .

Figure 8 plots the MSEs of a 40-step NLMS and RLS adaptive predictors for a stationary and a chirped input signal with chirp rate $\psi = 5\pi e - 4$, signal pole location $a = 0.99$, input signal power $P_s = 1$, $\text{SNR} = 20$ dB, and $M = 25$. For NLMS predictors, the MSEs obtained by the error transfer function approach and the simulation results are compared for both stationary and chirped inputs. The simulation results of the RLS adaptive predictor for stationary inputs reveal that the nonlinear effects are negligible for RLS algorithms. Comparing the results with Figure 7, the range of the adaptive filter step-size μ over which the NLMS adaptive predictors outperform the finite-length Wiener predictor is much larger, and the magnitude of the nonlinear effect is significant at optimal step-size (in this case, the optimal step-size for the adaptive predictor to achieve minimum MSE is about $\mu = 0.8$).

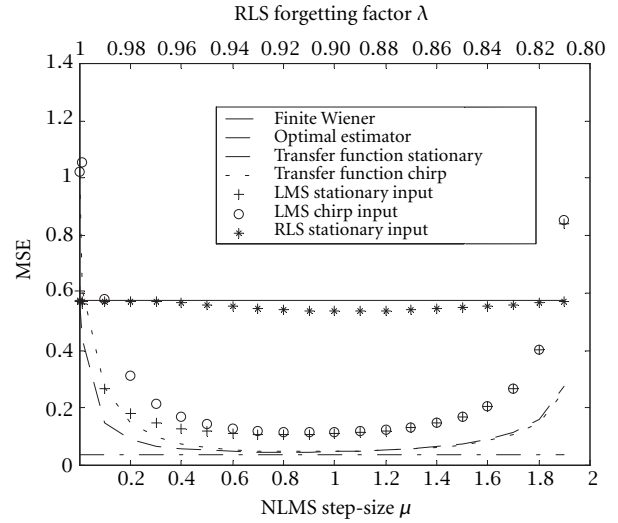


FIGURE 8: MSEs of NLMS and RLS 40-step predictors as a function of adaptation constant, with $\text{SNR} = 20$ dB, $M = 25$, $a = 0.99$, chirp rate $\psi = 5\pi e - 4$.

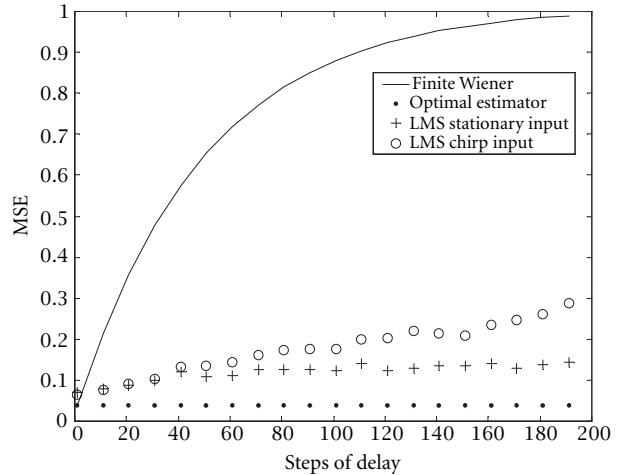


FIGURE 9: MSEs of NLMS predictor at optimal step-size as a function of prediction distance Δ , with $\text{SNR} = 20$ dB, $M = 25$, $a = 0.99$, chirp rate $\psi = 5\pi e - 4$.

One possible explanation for this is that for multiple-step prediction, the additional data which is available to adaptive predictors but not available to the finite-length Wiener predictor consists of two parts: $\{x_{n-1}^c, x_{n-2}^c, \dots, x_{n-(\Delta-1)}^c\}$ and $\{x_{n-(\Delta+M)}^c, x_{n-(\Delta+M+1)}^c, \dots, x_\infty^c\}$, and for one-step prediction, only the second part is available. The main contribution to the nonlinear effects is the first part and with the increase of prediction distance Δ , the correlation between the desired response x_n^c and the second part decreases, and the second part has less contribution to current estimation.

Figure 9 compares the MSEs of the finite-length Wiener predictor, the optimal estimator with the MSEs obtained in simulations achieved at optimal step-size μ_{opt} as a function

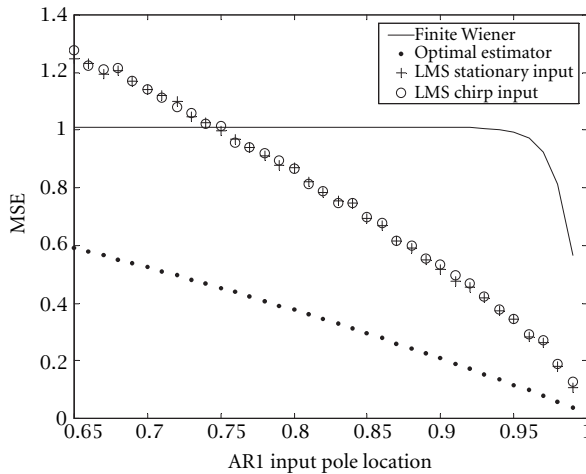


FIGURE 10: MSEs of 40-step NLMS adaptive predictor at optimal step-size as a function of input signal pole location a , with SNR = 20 dB, $M = 25$, chirp rate $\psi = 5\pi e - 4$.

of prediction distance Δ . It shows that with the above parameters, the LMS adaptive filter outperforms Wiener filter for $\Delta \geq 5$ and the nonlinear effect becomes more significant with increasing Δ .

Figure 10 is a plot of the various MSEs versus input signal pole location a for a 40-step predictor. It shows that the range of input signal pole location over which the nonlinear effect is observable is from about 0.75 to around 1. This range is also much larger compared to the one-step prediction case.

7. CONCLUSIONS

In conclusion, this paper shows that for very narrowband input signals, either stationary or nonstationary, traditional analysis using the independence assumptions is not valid and the nonlinear effect of the adaptive filter must be considered. For narrowband input signals embedded in AWGN, the LMS adaptive predictor can outperform the finite-length Wiener predictor in steady-state MSE. These cases arise when the adaptive filter uses more information than the finite-length Wiener filter. It shows that the nonlinear effect of one-step LMS adaptive predictors is small and only observable for a narrow range of input signal and adaptive filter parameters, and it is significant for multiple-step LMS adaptive predictors for a wide range of parameters. A transform is defined to convert the chirped input signal to baseband stationary input signal, and an error transfer function approach is derived for chirped input signals to approximate the total steady-state MSE of the LMS adaptive predictors. The performance of the one-step infinite-length Wiener predictor is used as the optimal estimator to bound the performance of adaptive Δ -step predictors. The nonlinear effects are much larger for the LMS adaptive predictor than for the exponentially weighted RLS predictor for the case examined.

ACKNOWLEDGEMENTS

This work was supported by the NSF Industry/University Cooperative Research Center on Ultra-High Speed Integrated Circuits and Systems (ICAS) at the University of California, San Diego.

We are grateful to A. A. (Louis) Beex for the very useful discussion and comments. Dr. Beex is a professor at the Bradley Department of Electrical and Computer Engineering at Virginia Polytechnic Institute and State University, USA.

REFERENCES

- [1] S. Haykin, *Adaptive Filter Theory*, Prentice Hall, New Jersey, 3rd edition, 1996.
- [2] J. R. Glover, "Adaptive noise cancelling applied to sinusoidal interferences," *IEEE Trans. Acoustics, Speech, and Signal Processing*, vol. ASSP-25, no. 6, pp. 484–491, 1977.
- [3] J. R. Zeidler, "Performance analysis of LMS adaptive prediction filters," in *Proceedings of the IEEE*, 1990, vol. 78, pp. 1781–1806.
- [4] M. Shensa, "Non-Wiener solutions for the adaptive noise canceller with a noisy reference," *IEEE Trans. Acoustics, Speech, and Signal Processing*, vol. 28, pp. 469–473, August 1980.
- [5] N. J. Bershad and P. L. Feintuch, "Non-Wiener solutions for the LMS algorithm—a time domain approach," *IEEE Trans. Signal Processing*, vol. SP-43, no. 5, pp. 1273–1275, 1995.
- [6] N. J. Bershad and J. C. M. Bermudez, "Sinusoidal interference rejection analysis of an LMS adaptive feedforward controller with a noisy periodic reference," *IEEE Trans. Signal Processing*, vol. 46, no. 5, pp. 1298–1313, 1998.
- [7] J. C. M. Bermudez and N. J. Bershad, "Non-Wiener behavior of the filtered LMS algorithm," *IEEE Trans. on Circuits and Systems II: Analog and Digital Signal Processing*, vol. 46, no. 8, pp. 1110–1114, 1999.
- [8] K. Quirk, L. Milstein, and J. R. Zeidler, "A performance bound for the LMS estimator," *IEEE Transactions on Information Theory*, vol. 46, no. 3, pp. 1150–1158, 2000.
- [9] M. Reuter and J. R. Zeidler, "Non-linear effects in LMS adaptive equalizers," *IEEE Trans. Signal Processing*, vol. 47, no. 6, pp. 1570–1579, 1999.
- [10] S. Barbarossa and R. Torti, "Chirped-OFDM for transmissions over time-varying channels with linear delay/Doppler spreading," in *Proceeding of the IEEE Conf. on Acoustics, Speech and Signal Processing*, May 2001.
- [11] 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, no. 2, pp. 363–376, 1997.
- [12] P. Wei, *Performance evaluation of adaptive filtering algorithms for the mobile communication environment*, Ph.D. thesis, University of California, San Diego, 1995.
- [13] O. Macchi and N. J. Bershad, "Adaptive recovery of a chirped sinusoid in noise, part 1: performance of the RLS algorithm," *IEEE Trans. Acoustics, Speech, and Signal Processing*, vol. ASSP-39, no. 3, pp. 583–594, 1991.
- [14] N. J. Bershad and O. Macchi, "Adaptive recovery of a chirped sinusoid in noise, part 2: performance of the LMS algorithm," *IEEE Trans. Acoustics, Speech, and Signal Processing*, vol. ASSP-39, no. 3, pp. 595–602, 1991.
- [15] O. Macchi, N. J. Bershad, and M. Mboup, "Steady-state superiority of LMS over LS for a time-varying line enhancer in a noisy environment," in *IEE Proceedings Part F-Radar and Signal Processing*, vol. 138, pp. 354–360, 1991.

Jun Han received the B.Eng. degree in Applied Electronic Technique from the University of Petroleum, 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 a Ph.D. candidate in the Department of Electrical and Computer Engineering at the University of California, San Diego, USA. His current research interests include digital signal processing in wireless communications and performance analysis of adaptive filters.



James R. Zeidler (IEEE 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 in the Electrical and Computer Engineering Department at the University of California, San Diego, CA, USA 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 received the B.S. degree (with Honors) in Electrical Engineering from the Moore School of Electrical Engineering, University of Pennsylvania, PA, and the M.S. and Ph.D. degrees in Electrical Engineering from the Polytechnic Institute of Brooklyn, NY, USA. 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 over the years 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 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 Raytheon Rockwell, and TRW as team members. Dr Ku is a member of Eta Kappa Nu, Tau Beta Pi, Sigma.

