Ability of Adaptive Filters to Track Carrier Offsets and Channel Nonstationarities

Nabil R. Yousef, *Student Member, IEEE,* and Ali H. Sayed, *Fellow, IEEE*

***Abstract*—This paper studies the tracking performance of adaptive filters operating in the presence of two sources of nonstationarities: carrier frequency offsets and random channel variations. Both impairments are common in digital communications due to mismatches between transmitter and receiver carrier generators and channel fading. The paper derives expressions for the mean-square error and shows how filter performance is degraded under such nonstationary conditions. Selections of step sizes for optimal tracking performance are derived, different adaptive algorithms are compared, and supporting simulation results are provided.**

***Index Terms*—Adaptive filter, carrier offsets, mean-square error, nonstationary environment, tracking analysis.**

I. Introduction

ADAPTIVE filters are often used in nonstationary environments where they are required to track time variations in an unknown system or channel [1]–[4]. The ability of adaptive filters to track such variations has received considerable attention in the literature over the last two decades (see, e.g., [5]–[10]). Most of the existing works, however, have focused on the case of random system nonstationarities whereby the channel is assumed to vary according to a random-walk or Markovian model (see, e.g., [10]).

A different scenario is very common in communication systems where mismatches between the transmitter and receiver carrier generators result in periodic system variations. These variations can be damaging to the performance of adaptive filters, even for very small carrier frequency offsets (see, e.g., [11], [12]). The ability of adaptive filtering algorithms to track such periodic system variations is not yet fully understood. A recent contribution in this regard is the work [11], which performed a first-order analysis of the performance of the least-mean-squares (LMS) algorithm in the presence of a carrier frequency offset only. Another earlier contribution is the work [9]. The effects of such carrier frequency offsets on other adaptive algorithms, i.e., other than LMS, have not been addressed in the literature. Furthermore, the *combined* effects of both cyclic and random system nonstationarities on the tracking performance of adaptive algorithms has remained largely an open issue, even for LMS.

Manuscript received November 3, 2000; revised March 26, 2002. This work was supported in part by the National Science Foundation under awards CCR-9732376 and ECS-9820765. The associate editor coordinating the review of this paper and approving it for publication was Dr. Rick S. Blum.

N. R. Yousef was with the Electrical Engineering Department, University of California, Los Angeles, CA 90095 USA. He is now with Broadcom Corporation, Irvine, CA 92619 USA.

A. H. Sayed is with the Electrical Engineering Department, University of California, Los Angeles, CA 90095 USA (e-mail: sayed@ee.ucla.edu).

Publisher Item Identifier S 1053-587X(02)05653-2.

The purpose of this paper is to present a framework for the tracking analysis of adaptive algorithms that handles simultaneously both cyclic and random system nonstationarities. In particular, the results will allow us to quantify the degradation in performance that results, for example, from carrier offsets. The results will also suggest optimal choices for filter parameters (e.g., step-sizes) in order to minimize the effect of such offsets on filter performance. Several supporting simulations are provided.

In Section II, we motivate the data model that is employed in Sections III–VI.

A. Model

Let $\{u(i)\}$ denote a sequence that is transmitted over an unknown channel of finite impulse response $\mathbf{w}_i^o$ of order M. It is assumed that the channel varies in time according to the rule

$$\mathbf{w}_i^o = \mathbf{w}^o + \theta_i \tag{1}$$

where $\mathbf{w}^o$ is a constant vector, and θ_i is a random perturbation. For example, in the case of a fading channel, $\mathbf{w}^o$ would represent the nonfading part of the channel, whereas θ_i would represent the fading part. In addition, such FIR models are suitable for modeling multipath components.

In general, we can be more specific about the behavior of θ_i. For example, in the case of a fading channel again, the perturbation θ_i can be modeled to a reasonable extent as an autoregressive (AR) process of some order p [13]. A widely used approximation of this process is the AR(1) model, which corresponds to $p = 1$. Furthermore, if the channel components are assumed to fade independently and following the same statistical model, the process can be modeled by the following AR(1) model:

$$\boldsymbol{\theta}_{i+1} = \alpha\boldsymbol{\theta}_i + \mathbf{q_i} \tag{2}$$

where

$$0 \leq |\alpha| < 1 \tag{3}$$

and $\mathbf{q}_i$ is a zero-mean stationary random vector process with a positive-definite covariance matrix $\mathbf{Q} = E(\mathbf{q}_i\mathbf{q}_i^*)$. For a Rayleigh fading channel of maximum Doppler frequency f_D, one has $\alpha = J_o(2\pi f_D T)$ and $\mathbf{Q} = \sqrt{1-\alpha^2}\mathbf{I}$, where

- J_o Bessel function of first kind and order zero;
- T sampling period of the digital communication system;
- $\mathbf{I}$ identity matrix (see, e.g., [14]).

If we denote the regressor of the channel by $\mathbf{u}_i$

$$\mathbf{u}_i = [\,u(i) \quad u(i-1) \quad \ldots \quad u(i-M+1)\,]$$

1053-587X/02$17.00 © 2002 IEEE

then the undisturbed output of the channel is given by the inner product $\mathbf{u}_i\mathbf{w}_i^o$. However, due to measurement noise and mismatches in carrier frequencies at the transmitter and receiver, the received data is actually modeled by

$$d(i) = \mathbf{u}_i\mathbf{w}_i^o e^{j\Omega i} + v(i) \tag{4}$$

where $v(i)$ is measurement noise, and the multiplicative term $e^{j\Omega i}$ accounts for the carrier offset [11], [12].[1]

The term $e^{j\Omega i}$ could also be used to model Doppler channel variations in a wireless scenario, which result from reflections of the transmitted signal off a remote object moving with constant speed [such as a low flying airplane (airplane flutter) or a swaying tower or skyscraper [15]]. Actually, many digital communication standards use the ability of digital communication receivers to track such Doppler shifts as a performance index for their ability to track time-varying channels (see [15] for an equalization example in the context of terrestrial television channels).

In summary, the above discussion motivates us to focus in this paper on data $\{d(i), \mathbf{u}_i\}$ that arise from a model of the form

$$\begin{aligned}
\mathbf{w}_i^o &= \mathbf{w}^o + \theta_i && (5)\\
\boldsymbol{\theta}_{i+1} &= \alpha\, \boldsymbol{\theta}_i + \mathbf{q}_i && (6)\\
d(i) &= \mathbf{u}_i\mathbf{w}_i^o e^{j\Omega i} + v(i) && (7)\\
& 0 \le |\alpha| < 1 &&\\
E\left(\mathbf{q}_i\mathbf{q}_j^*\right) &= \mathbf{Q}\delta_{ij}. &&
\end{aligned}$$

This model includes the effects of both cyclic and random system nonstationarities (through Ω and $\mathbf{q}_i$), both of which are common impairments in communication systems and especially in applications that involve channel estimation, channel equalization, and intersymbol-interference cancellation.

B. Adaptive Filtering Algorithms

As stated previously, the purpose of this paper is to study the ability of LMS-type adaptive filters to estimate and track such cyclic and random variations in $\mathbf{w}_i^o$. The LMS family of algorithms is the most widely used in digital communications applications due to its simplicity and stability properties. We therefore consider general adaptive schemes of the form

$$\mathbf{w}_{i+1} = \mathbf{w}_i + \mu\mathbf{u}_i^* f_e(i) \tag{8}$$

where "$*$" denotes Hermitian conjugation (complex conjugation for scalars), $\mathbf{w}_i$ is the estimate for $\mathbf{w}_i^o$ at iteration i, μ is the step size, and $f_e(i)$ is the generic scalar function of the output estimation error

$$e(i) = d(i) - \mathbf{u}_i\mathbf{w}_i.$$

Different choices for $f_e(i)$ result in different adaptive algorithms. Table I defines $f_e(i)$ for some famous special cases of (8); see [1], [2].[2]

[1]Here, we have neglected carrier phase noise for simplicity. Furthermore, the offset frequency Ω is assumed to be constant over time. A more general model would include a time-varying term of the form $e^{j(\Omega i+\phi(i))}$ to account for offset frequency and phase noise term.

[2]The list in the table assumes real-valued data. For complex-valued data, we replace e^3 by $e|e|^2$, and define sign$[a+jb]$ by $(1/\sqrt{2})(\text{sign}[a] + j\text{sign}[b])$.

TABLE I
EXAMPLES FOR $f_e(i)$

ALGORITHM	$f_e(i)$
LMS	$e(i)$
LMF	$e^3(i)$
LMMN	$\delta e(i) + (1-\delta)e^3(i)$
SA	$\text{sign}[e(i)]$
NLMS	$e(i)/\Vert\mathbf{u}_i\Vert^2$

An important performance measure for an adaptive filter is its steady-state mean-square-error (MSE), which is defined as

$$\text{MSE} = \lim_{i\to\infty} E|e(i)|^2 = \lim_{i\to\infty} E|v(i) + \mathbf{u}_i\tilde{\mathbf{w}}_i|^2$$

where the weight error vector $\tilde{\mathbf{w}}_i$ is defined by

$$\boxed{\tilde{\mathbf{w}}_i = \mathbf{w}_i^o e^{j\Omega i} - \mathbf{w}_i.} \tag{9}$$

Under the following often-realistic assumption:

A.1 The noise sequence $\{v(i)\}$ is iid, with variance σ_v^2, and statistically independent of the regressor sequence $\{\mathbf{u}_i\}$ (see, e.g., [1]–[4]), we find that the MSE is equivalently given by

$$\text{MSE} = \sigma_v^2 + \lim_{i\to\infty} E|\mathbf{u}_i\tilde{\mathbf{w}}_i|^2. \tag{10}$$

In the sequel, we proceed to derive expressions for the steady-state excess mean-square-error (EMSE)

$$\zeta = \lim_{i\to\infty} E|\mathbf{u}_i\tilde{\mathbf{w}}_i|^2 = \lim_{i\to\infty} E|e_a(i)|^2$$

for various algorithms, along with values for the optimum algorithm parameters that minimize the EMSE. By deriving these expressions, we arrive at several results on the tracking performance of adaptive filters. These results help clarify the effect of cyclic nonstationarities on the algorithms of Table I. With the exception of a first-order analysis for LMS in [11], which was performed in the absence of random nonstationarities (i.e., for $\alpha = 0$ and $\mathbf{q}_i = 0$) and a related analysis in [9], such effects are not yet fully understood. Furthermore, while common tracking analysis in the literature assume a random-walk model with $\Omega = 0$, $\mathbf{w}^o = \mathbf{0}$ and $\alpha = 1$ (see, e.g., [16]–[18]), it turns out that assuming $\alpha = 1$ when cyclic nonstationarities are present does not lead to practical design expressions.

II. FUNDAMENTAL ENERGY RELATION

In this section, we derive an energy conservation relation and explain its relevance to mean-square analysis. Thus, using (7)–(9), we obtain the following recursion for the weight-error vector:

$$\tilde{\mathbf{w}}_{i+1} = \tilde{\mathbf{w}}_i - \mu\mathbf{u}_i^* f_e(i) + \mathbf{c}_i e^{j\Omega i} \tag{11}$$

where

$$\mathbf{c}_i \triangleq \mathbf{w}^o(e^{j\Omega} - 1) + \boldsymbol{\theta}_i(\alpha e^{j\Omega} - 1) + \mathbf{q}_i e^{j\Omega}. \tag{12}$$

We further define *a priori* and *a posteriori* estimation errors as

$$\boxed{e_a(i) = \mathbf{u}_i\tilde{\mathbf{w}}_i, \quad e_p(i) = \mathbf{u}_i\left(\tilde{\mathbf{w}}_{i+1} - \mathbf{c}_i\mathbf{e}^{j\Omega i}\right)}$$

and use the data model (7) to find that

$$e(i) = e_a(i) + v(i).$$

Moreover, if we multiply (11) by $\mathbf{u}_i$ from the left, we also find that

$$e_p(i) = e_a(i) - \mu\|\mathbf{u}_i\|^2 f_e(i). \tag{13}$$

Substituting (13) into (11), we obtain, for nonzero $\mathbf{u}_i$

$$\tilde{\mathbf{w}}_{i+1} = \tilde{\mathbf{w}}_i - \frac{\mathbf{u}_i^*}{\|\mathbf{u}_i\|^2}[e_a(i) - e_p(i)] + \mathbf{c}_i e^{j\Omega i}. \tag{14}$$

By evaluating the energies of both sides of this equation, we find that

$$\|\tilde{\mathbf{w}}_{i+1} - \mathbf{c}_i\mathbf{e}^{j\Omega i}\|^2 + \frac{1}{\|\mathbf{u}_i\|^2}|e_a(i)|^2 = \|\tilde{\mathbf{w}}_i\|^2 + \frac{1}{\|\mathbf{u}_i\|^2}|e_p(i)|^2. \tag{15}$$

When $\mathbf{u}_i = 0$, it is obviously true that

$$\|\tilde{\mathbf{w}}_{i+1} - \mathbf{c}_i\mathbf{e}^{j\Omega i}\|^2 = \|\tilde{\mathbf{w}}_i\|^2. \tag{16}$$

Both results (15) and (16) can be grouped together into a single equation by defining

$$\bar{\mu}(i) = \left(\|\mathbf{u}_i\|^2\right)^\dagger$$

in terms of the pseudo-inverse of a scalar so that we obtain

$$\boxed{\|\tilde{\mathbf{w}}_{i+1} - \mathbf{c}_i\mathbf{e}^{j\Omega i}\|^2 + \bar{\mu}(i)|e_a(i)|^2 = \|\tilde{\mathbf{w}}_i\|^2 + \bar{\mu}(i)|e_p(i)|^2.} \tag{17}$$

This energy-conservation relation, which was first noted in [19]–[21], holds for *all* adaptive algorithms whose recursions are of the form given by (8); there are no approximations involved. It shows how the energies of the weight error vectors at two successive time instants are related to the energies of the *a priori* and *a posteriori* estimation errors.

Relation (17) was used in [22] in the special case $\Omega = 0$, $\alpha = 1$ to study the mean-square-error performance of adaptive filters. Some care is required to extend the analysis to the context of this paper due to the complications introduced by the *simultaneous* presence of cyclic and random nonstationarities.

A. Relevance to the Tracking Analysis

We are interested in using the energy relation (17) to evaluate the EMSE of an adaptive filter once it reaches steady state. Thus, using (11)–(13) and $E\|\tilde{\mathbf{w}}_{i+1}\|^2 = E\|\tilde{\mathbf{w}}_i\|^2$ in steady state, and taking expectations of both sides, it is shown in Appendix A that (17) becomes

$$\begin{aligned}2\mu \mathrm{Re} E\left(e_a^*(i)f_e(i)\right) =& \mu^2 E\left(\|\mathbf{u}_i\|^2\left|f_e(i)\right|^2\right) + \mathrm{Tr}(\mathbf{Q})\\ &+ |1-e^{j\Omega}|^2\mathrm{Tr}(\mathbf{W}^o)\\ &+ |1-\alpha e^{j\Omega}|^2\mathrm{Tr}(\mathbf{\Theta})\\ &- 2\mathrm{Re}\Bigg[\left(1-e^{-j\Omega}\right)\mathbf{w}^{o*}\\ &\qquad \times E\left(\left(\tilde{\mathbf{w}}_i - \mu\mathbf{u}_i^* f_e(i)\right)e^{-j\Omega i}\right)\Bigg]\\ &- 2\mathrm{Re}\Bigg[\left(1-\alpha^* e^{-j\Omega}\right)\\ &\qquad \times E\left(\boldsymbol{\theta}_i^*\left(\tilde{\mathbf{w}}_i - \mu\mathbf{u}_i^* f_e(i)\right)e^{-j\Omega i}\right)\Bigg]\end{aligned} \tag{18}$$

where

$$\mathbf{W}^o \triangleq \mathbf{w}^o\mathbf{w}^{o*}, \quad \mathbf{\Theta} \triangleq \lim_{i\to\infty} E\left(\boldsymbol{\theta}_i\boldsymbol{\theta}_i^*\right). \tag{19}$$

It is easy to verify from (5) and (6) that

$$\mathbf{\Theta} = \frac{\mathbf{Q}}{1-|\alpha|^2}. \tag{20}$$

Equation (18) can now be solved for the steady-state EMSE of various adaptive algorithms from Table I. This requires that several terms in (18) be evaluated and the resulting expression solved for

$$\boxed{\zeta = \lim_{i\to\infty} E\left|e_a(i)\right|^2.}$$

Due to space limitations, we illustrate the procedure for LMS and list the results for the other algorithms; see also [18], [22], and [23] for other steady-state and tracking results in the absence of cyclic nonstationarities.

III. Tracking Analysis

A. LMS *Algorithm*

For LMS, we have

$$f_e(i) = e(i) = e_a(i) + v(i). \tag{21}$$

In order to proceed, we need to evaluate the terms

$$E\left(\tilde{\mathbf{w}}_i\right) \text{ and } E\left(\tilde{\mathbf{w}}_i\boldsymbol{\theta}_i^*\right)$$

which appear in (18). We start with the first term. It turns out that $E(\tilde{\mathbf{w}}_i)$ takes the generic form $\mathbf{v}\mathbf{e}^{j\Omega i}$ in steady state for some $\mathbf{v}$. To verify this result, we call on the following steady-state independence assumption:[3]

A.2 At steady state, $\tilde{\mathbf{w}}_i$ is statistically independent of $\mathbf{u}_i$.[4]

Lemma 1: Consider the LMS recursion

$$\mathbf{w}_{i+1} = \mathbf{w}_i + \mu\mathbf{u}_i^* e(i)$$

where $e(i) = d(i) - \mathbf{u}_i\mathbf{w}_i$, and the data $\{d(i), \mathbf{u}_i\}$ is assumed to satisfy the model (5)–(7). In steady state (i.e., $i \to \infty$), it holds that

$$\boxed{E\left(\tilde{\mathbf{w}}_i\right) = \mathbf{v}\mathbf{e}^{j\Omega i}}$$

where

$$\boxed{\mathbf{v} = \left[\mathbf{I} - \mu\mathbf{\Gamma} - e^{j\Omega}\mathbf{I}\right]^{-1}\mathbf{w}^o\left(1 - e^{j\Omega}\right)} \tag{22}$$

and $\mathbf{\Gamma} = \mathbf{R} \triangleq E\mathbf{u}_i\mathbf{u}_i^*$.

Proof: Let $\mathbf{v}_i = E\tilde{\mathbf{w}}_i$. Applying the expectation operator to both sides of (11) and using (12), (21), A.1, and A.2, we obtain

$$\mathbf{v}_{i+1} = (\mathbf{I} - \mu\mathbf{R})\mathbf{v}_i + \mathbf{w}^o(e^{j\Omega} - 1)e^{j\Omega i}. \tag{23}$$

To proceed we introduce the eigenvalue decomposition of the matrix $\mathbf{R}$

$$\mathbf{R} = \mathbf{U}^*\mathbf{\Lambda}\mathbf{U} \tag{24}$$

[3]We are only requiring $\tilde{\mathbf{w}}_i$ to be independent of $\mathbf{u}_i$ in steady state. This is a weaker assumption than the usual full-blown independence assumptions [2].

[4]Of course, $\tilde{\mathbf{w}}_i$ is not statistically independent of $\mathbf{u}_i$, except in very special cases. However, this assumption is realistic for small step-size μ, as well as for long filters. Intuitively, the update term in (8) is relatively small for small step sizes, and the statistical dependence of $\tilde{\mathbf{w}}_i$ on $\mathbf{u}_i$ becomes weak. Furthermore, in steady state, the error $e(i)$ is also small, which makes the update term in (8) even smaller.

where $\mathbf{U}$ is a unitary matrix, and $\mathbf{\Lambda}$ is a positive-definite diagonal matrix with entries

$$\mathbf{\Lambda} = \text{diag}\{\lambda_1, \lambda_2, \ldots, \lambda_M\} \tag{25}$$

where λ_n is the nth eigenvalue of $\mathbf{R}$. Multiplying both sides of (23) by $\mathbf{U}$ from the left and using (24), we get

$$\mathbf{v}'_{i+1} = (\mathbf{I} - \mu\mathbf{\Lambda})\,\mathbf{v}'_i + \mathbf{c}' e^{j\Omega i}$$

where

$$\mathbf{v}'_i \triangleq \mathbf{U}\mathbf{v}_i, \quad \mathbf{c}' \triangleq \mathbf{U}\mathbf{w}^o(e^{j\Omega} - 1).$$

We can now write the following recurrence relation for the nth element of the vector $\mathbf{v}'_i$:

$$v'_n(i+1) = (1 - \mu\lambda_n)\,v'_n(i) + c'_n e^{j\Omega i}, \quad n = 1, \ldots, M.$$

Taking the z transform of both sides of this equation, we get

$$z v'_n(z) = (1 - \mu\lambda_n)\,v'_n(z) + c'_n \frac{z}{z - e^{j\Omega}}.$$

Using partial fractions, we can rewrite the above equation as

$$v'_n(z) = \frac{a_n}{z - (1 - \mu\lambda_n)} + \frac{b_n}{z - e^{j\Omega}}$$

where a_n and b_n are constants. Obtaining the inverse z transforms of both sides leads to

$$v'_n(i) = a'_n\,(1 - \mu\lambda_n)^i + b'_n e^{j\Omega i}$$

where

$$a'_n = a_n e^{-j\Omega}, \quad b'_n = b_n e^{-j\Omega}.$$

We now impose the condition (which is actually necessary for algorithm convergence)

$$\mu < \frac{2}{\lambda_{\max}} \tag{26}$$

where $\lambda_{\max} = \max[\lambda_1, \lambda_2, \ldots, \lambda_M]$. Then, we get that in steady state, $(i \to \infty)$, $v'_n(i)$ is given by

$$v'_n(i) = b'_n e^{j\Omega i}.$$

By using $\mathbf{v}'_i = \mathbf{U}\mathbf{v}_i$, we conclude that the vector $\mathbf{v}_i$ converges, in steady state, to

$$\mathbf{v}_i = \mathbf{v} e^{j\Omega i} \tag{27}$$

for some time-independent vector $\mathbf{v}$. To calculate $\mathbf{v}$, we substitute (27) into (23) to get (22). ■

In a similar vein, we now verify that the matrix

$$\mathbf{W}_i \triangleq E\left(\tilde{\mathbf{w}}_i \boldsymbol{\theta}_i^*\right)$$

takes the form $\mathbf{W}\mathbf{e}^{j\Omega i}$ in steady state for some matrix $\mathbf{W}$.

Lemma 2: Consider the same setting of Lemma 1. It then holds in steady-state that

$$\boxed{E(\mathbf{W}_i) = \mathbf{W}\mathbf{e}^{j\Omega i}} \tag{28}$$

where

$$\boxed{\mathbf{W} = \left[\alpha^*(\mathbf{I} - \mu\mathbf{\Gamma}) - e^{j\Omega}\mathbf{I}\right]^{-1}\mathbf{C}} \tag{29}$$

and

$$\boxed{\mathbf{C} \triangleq \alpha^*\left(1 - \alpha e^{j\Omega}\right)\mathbf{\Theta} - e^{j\Omega}\mathbf{Q}.} \tag{30}$$

Proof: If we substitute (21) into (11), multiply by $\boldsymbol{\theta}_{i+1}^* e^{-j\Omega i}$ from the right, and apply the expectation operator to both of its sides, we get

$$\mathbf{W}_{i+1} = \alpha^*(\mathbf{I} - \mu\mathbf{R})\,\mathbf{W}_i - \mathbf{C}e^{j\Omega i}. \tag{31}$$

Multiplying both sides by $\mathbf{U}$ from the left and using (24), we get

$$\mathbf{W}'_{i+1} = \alpha^*(\mathbf{I} - \mu\mathbf{\Lambda})\,\mathbf{W}'_i - \mathbf{C}'e^{j\Omega i}$$

where

$$\mathbf{W}'_i \triangleq \mathbf{U}\mathbf{W}_i \quad \mathbf{C}' \triangleq \mathbf{U}\mathbf{C}.$$

Using similar arguments to the ones that led to (27), it is straightforward to verify that each element of the matrix $\mathbf{W}_i$ converges to a constant times the time-varying exponent $\mathbf{e}^{j\Omega i}$ when the two conditions given by (3) and (26) hold. It then follows that (28) holds. To evaluate $\mathbf{W}$, we use (31) to obtain

$$e^{j\Omega}\mathbf{W} = \alpha^*(\mathbf{I} - \mu\mathbf{R})\,\mathbf{W} - \mathbf{C}.$$

Solving for $\mathbf{W}$ yields (29). ■

With expressions for both $\mathbf{v}$ and $\mathbf{W}$ in hand, we can now proceed to solve (18) for the EMSE of LMS. Substituting (21) into (18) and using A.1, we obtain

$$\begin{aligned}
2\mu\zeta^{\mathsf{LMS}} =& \mu^2\sigma_v^2 \text{Tr}(\mathbf{R}) + \mu^2 E\left(\|\mathbf{u}_i\|^2\,|e_a(i)|^2\right) + \text{Tr}(\mathbf{Q}) \\
&+ |1 - e^{j\Omega}|^2 \text{Tr}(\mathbf{W}^o) + |1 - \alpha e^{j\Omega}|^2 \text{Tr}(\mathbf{\Theta}) \\
&- 2\text{Re}\left[\left(1 - e^{-j\Omega}\right)\mathbf{w}^{o*}(\mathbf{I} - \mu\mathbf{R})\,\mathbf{v}\right] \\
&- 2\text{Re Tr}\left[\left(1 - \alpha^* e^{-j\Omega}\right)(\mathbf{I} - \mu\mathbf{R})\,\mathbf{W}\right].
\end{aligned} \tag{32}$$

To solve for ζ^{LMS}, we consider three typical cases.

1) For sufficiently small μ, we can assume that the term $\mu^2 E\left(\|\mathbf{u}_i\|^2 |e_a(i)|^2\right)$ is negligible so that

$$\boxed{\zeta^{\mathsf{LMS}} = \tfrac{\mu}{2}\sigma_v^2 \text{Tr}(\mathbf{R}) + \tfrac{\mu^{-1}}{2}\beta \quad (\text{small } \mu)} \tag{33}$$

where

$$\begin{aligned}
\beta =& |1 - e^{j\Omega}|^2 \text{Re Tr}\left[\mathbf{W}^o(\mathbf{I} - 2\mathbf{X})\right] \\
&+ |1 - \alpha e^{j\Omega}|^2 \text{Re Tr}\left[\mathbf{\Theta}(\mathbf{I} - 2\alpha^*\mathbf{X}_\alpha)\right] \\
&+ \text{Re Tr}\left[\mathbf{Q}\left(\mathbf{I} - 2\left(\alpha^* - e^{j\Omega}\right)\mathbf{X}_\alpha\right)\right]
\end{aligned} \tag{34}$$

and $\mathbf{X}$ and $\mathbf{X}_\alpha$ are defined by

$$\begin{aligned}
\mathbf{X} &\triangleq (\mathbf{I} - \mu\mathbf{\Gamma})\left(\mathbf{I} - \mu\mathbf{\Gamma} - e^{j\Omega}\mathbf{I}\right)^{-1} \\
\mathbf{X}_\alpha &\triangleq (\mathbf{I} - \mu\mathbf{\Gamma})\left[\alpha^*(\mathbf{I} - \mu\mathbf{\Gamma}) - e^{j\Omega}\mathbf{I}\right]^{-1}.
\end{aligned} \tag{35}$$

2) For larger values of μ, (32) can be solved by imposing the following assumption (which is realistic for longer filter).

A.3 At steady state, $\mu\|\mathbf{u}_i\|^2$ is statistically independent of $|e_a(i)|^2$.

This assumption is not needed in the case of constant-modulus data that arises in some adaptive filtering applications (see, e.g., [25]). Using A.3 and (32), we obtain

$$\boxed{\zeta^{\mathsf{LMS}} = \frac{\mu\sigma_v^2\mathrm{Tr}(\mathbf{R})+\mu^{-1}\beta}{2-\mu\mathrm{Tr}(\mathbf{R})} \quad (\text{large } \mu)} \tag{36}$$

3) For Gaussian white-input signals ($\mathbf{R}=\sigma_u^2\mathbf{I}$), (32) can be more accurately solved by using A.2 to yield

$$\boxed{\zeta^{\mathsf{LMS}} = \frac{\mu M\sigma_u^2\sigma_v^2+\mu^{-1}\beta}{2-\mu(M+\lambda)\sigma_u^2} \quad (\text{Gaussian})} \tag{37}$$

where M is the filter length, $\lambda = 1$ if the $\{\mathbf{u}_i\}$ are complex-valued, and $\lambda = 2$ if the $\{\mathbf{u}_i\}$ are real valued. Moreover, β is now given by

$$\begin{aligned}\beta =& \mu\sigma_u^2\left(2-\mu\sigma_u^2\right)\frac{|1-e^{j\Omega}|^2}{|1-\mu\sigma_u^2-e^{j\Omega}|^2}\|\mathbf{w}^o\|^2\\ &+\left(1-|\alpha|^2\left(1-\mu\sigma_u^2\right)^2\right)\frac{|1-\alpha e^{j\Omega}|^2}{|\alpha^*\left(1-\mu\sigma_u^2\right)-e^{j\Omega}|^2}\mathrm{Tr}\left(\mathbf{\Theta}\right)\\ &+\mathrm{Re}\left[1-\frac{2\left(\alpha^*-e^{j\Omega}\right)\left(1-\mu\sigma_u^2\right)}{\alpha^*\left(1-\mu\sigma_u^2\right)-e^{j\Omega}}\right]\mathrm{Tr}(\mathbf{Q}).\end{aligned} \tag{38}$$

To further understand the effect of the different types of system nonstationarities on the tracking performance of the LMS algorithm, we specialize the above results to the following two cases. For simplicity, we consider small values of μ and white input signals.

Carrier Offset Only $[Tr(\mathbf{Q})=0]$*:* In this case, the second and third terms on the right-hand side of (34) and (38) are equal to zero. Furthermore, for small values of Ω and $\mu\sigma_u^2 \gg (1-\cos\Omega)$, which is usually valid in practical cases, (38) can be approximated by

$$\beta \approx \frac{\Omega^2\left(2-\mu\sigma_u^2\right)}{\mu\sigma_u^2}\|\mathbf{w}^o\|^2. \tag{39}$$

Here, it can be seen that unlike the stationary case, the steady-state EMSE is not a monotonically increasing function of the step size μ. The EMSE is composed of two terms. The first term increases with μ, the noise variance σ_v^2, and σ_u^2. The second term decreases with μ and increases with the frequency offset Ω. This term becomes dominant for small values of μ and causes the EMSE to increase with the order of μ^2 when decreasing μ. Furthermore, it is clear that there exists a value of the algorithm step-size (μ_o) that minimizes the EMSE. This optimal value can be obtained by minimizing the EMSE[5] in (33) over μ, i.e., finding μ_o is equivalent to solving for the positive root of[6]

$$a_1\mu_o^3+a_2\mu_o-a_3=0 \tag{40}$$

where

$$a_1 = M\sigma_u^2\sigma_v^2, \quad a_2=\Omega^2\|\mathbf{w}^o\|^2, \quad a_3=\frac{4\Omega^2\|\mathbf{w}^o\|^2}{\sigma_u^2}.$$

A *rough* estimate for μ_o in this case is given by

$$\mu_o^{\mathsf{LMS}} \approx \left(\frac{4\Omega^2\|\mathbf{w}^o\|^2}{M\sigma_v^2\sigma_u^4}\right)^{1/3}. \tag{41}$$

Here, we can see that the optimum step size increases with the frequency offset Ω and with $\|\mathbf{w}^o\|^2$ and decreases with the noise variance σ_v^2 and the filter length M. Furthermore, the corresponding minimum achievable EMSE is approximately given by

$$\zeta_{\min}^{\mathsf{LMS}} = \frac{3}{4}\left(4M^2\sigma_v^4\sigma_u^2\Omega^2\|\mathbf{w}^o\|^2\right)^{1/3}. \tag{42}$$

Random Nonstationarity Only $[\Omega=0,\ \alpha=1]$*:* In this case, (34) reduces to

$$\beta=\mathrm{Tr}(\mathbf{Q}).$$

Here, the second term of the EMSE decreases with μ and increases with the random nonstationarity term Tr($\mathbf{Q}$). The optimum value of the step size in this case is given by

$$\mu_o^{\mathsf{LMS}} = \frac{1}{\sigma_v}\sqrt{\frac{\mathrm{Tr}(\mathbf{Q})}{\mathrm{Tr}(\mathbf{R})}}$$

which is the same expression given in [10].

For the more general case, the optimal value of μ can obtained by minimizing the EMSE, which is given by (33), over a dense grid of all possible values of μ.

B. LMF*,* LMNN*, sign, and* NLMS *Algorithms*

We now extend the results to the other adaptive filters in Table I.

It can be verified that expressions (22) and (29) still hold for the LMF, LMMN, sign, and NLMS algorithms. What changes is the value of the matrix $\mathbf{\Gamma}$ and the condition on μ for each algorithm. These values are listed in Table II, where

$$\begin{aligned}\gamma =& \delta+3\bar{\delta}\sigma_v^2, \quad \bar{\delta}=1-\delta \quad (\text{for LMMN})\\ \eta =& \sqrt{\frac{2}{\pi(\sigma_v^2+\zeta)}} \quad (\text{for sign-LMS}), \quad \mathbf{R}_N \triangleq E\left(\frac{\mathbf{u}_i^*\mathbf{u}_i}{\|\mathbf{u}_i\|^2}\right)\end{aligned}$$

and $\lambda_{\max}(\mathbf{R})$ and $\lambda_{\max}(\mathbf{R}_N)$ are the maximum eigenvalues of $\mathbf{R}$ and $\mathbf{R}_N$, respectively.

Substituting the expressions for $\mathbf{v}$ and $\mathbf{W}$ into (18) and following the same steps used for the LMS algorithm, we obtain in Appendix B the EMSE expressions that are listed in Table III.

[5]In general, the optimal step size should be obtained by minimizing the MSE expressions. Here, we minimize the MSE expression for small step sizes for simplicity.

[6]In general, the value of Ω is unknown. Still, the value of μ could be chosen based on an estimate for Ω or, in a worst-case design, based on the maximum expected value for Ω. The results of this paper indicate, among other things, how the performance of an adaptive filter is affected by Ω.

TABLE II
VALUES OF Γ AND CONDITIONS ON μ

ALGORITHM	Γ	μ
LMS	$\mathbf{R}$	$\mu < \frac{2}{\lambda_{max}(\mathbf{R})}$
LMF	$3\sigma_v^2\mathbf{R}$	$\mu < \frac{2}{3\sigma_v^2\lambda_{max}(\mathbf{R})}$
LMMN	$\gamma\mathbf{R}$	$\mu < \frac{1}{\gamma\lambda_{max}(\mathbf{R})}$
SA	$\eta\mathbf{R}$	$\mu < \frac{2}{\eta\lambda_{max}(\mathbf{R})}$
NLMS	$\mathbf{R}_N$	$\mu < \frac{2}{\lambda_{max}(\mathbf{R}_N)}$

TABLE III
EXPRESSIONS FOR THE MSE FOR VARIOUS ADAPTIVE FILTERS

ALGORITHM	EMSE
LMS (small μ)	$\frac{\mu}{2}\sigma_v^2\,\mathrm{Tr}(\mathbf{R}) + \frac{\mu^{-1}}{2}\beta$
LMS (larger μ)	$\frac{\mu\sigma_v^2\,\mathrm{Tr}(\mathbf{R})+\mu^{-1}\beta}{2-\mu\,\mathrm{Tr}(\mathbf{R})}$
LMS (Gaussian)	$\frac{\mu M\sigma_u^2\sigma_v^2+\mu^{-1}\beta}{2-\mu(M+\lambda)\sigma_u^2}$
LMMN (small μ)	$\frac{\mu a\,\mathrm{Tr}(\mathbf{R})+\mu^{-1}\beta}{2\gamma}$
LMMN (larger μ)	$\frac{\mu a\,\mathrm{Tr}(\mathbf{R})+\mu^{-1}\beta}{2\gamma-\mu b\,\mathrm{Tr}(\mathbf{R})}$
LMMN (Gaussian)	$\frac{\mu a M\sigma_u^2+\mu^{-1}\beta}{2\gamma-\mu b(M+\lambda)\sigma_u^2}$
SA	$\frac{\mu\,\mathrm{Tr}(\mathbf{R})+\mu^{-1}\beta}{2\eta}$
NLMS	$\frac{\mu\sigma_v^2+\mu^{-1}\beta/\,\mathrm{E}\left(\frac{1}{\lVert\mathbf{u}_i\rVert^2}\right)}{2-\mu}$

In Table III, β is as in (34),[7]

$$a = \left(\delta^2\sigma_v^2 + 2\delta\bar{\delta}\xi_v^4 + \bar{\delta}^2\xi_v^6\right), \quad b = \left(\delta^2 + 12\delta\bar{\delta}\sigma_v^2 + 15\bar{\delta}^2\xi_v^4\right)$$

with $\xi_v^4 = E|v(i)|^4$ and $\xi_v^6 = E|v(i)|^6$. Note that expressions for the LMF algorithm can directly be obtained by setting $\delta = 0$ in the corresponding LMMN expressions.

Carrier Offset Only $[Tr(\mathbf{Q}) = 0]$*:* In this case, expressions for the EMSE can be found by setting $\mathrm{Tr}(\mathbf{Q}) = 0$ in the expressions given in Table III. Furthermore, using the same procedure used for the LMS algorithm, the value of the algorithm step size (μ_o) that minimizes the EMSE can be found by solving for the positive root of

$$a_1\mu_o^3 + a_2\mu_o - a_3 = 0$$

where $a_2 = \Omega^2\|\mathbf{w}^o\|^2$ and the values of a_1, a_2, and a_3 for each algorithm are given in Table IV, along with a *rough* estimate for μ_o and the corresponding value of the minimum achievable EMSE $(\zeta_{\min})$; for the sign algorithm, we use the approximation $\eta \approx \sqrt{2/(\pi\sigma_v^2)}$.

Here, we can see that the optimum step size for each algorithm increases with the frequency offset Ω and with $\|\mathbf{w}^o\|^2$ and decreases with the noise variance σ_v^2 and the filter length M. In addition, note that the expression for the EMSE of the NLMS (which is given in Table III) can be minimized over μ to arrive at a value for the optimum step size.

[7]For complex-valued data, we replace γ with $\gamma' = \delta + 2\bar{\delta}\sigma_v^2$ and b by $b' = (\delta^2 + 8\delta\bar{\delta}\sigma_v^2 + 9\bar{\delta}^2\xi_v^4)$.

IV. COMPARISONS WITH LMS

A. LMMN *and* LMF

We now compare the ability of the LMF and LMMN algorithms to track variations in nonstationary environments with that of the LMS algorithm, which is known to have excellent tracking properties (see, e.g., [1], [2], and [10]). We focus only on the cyclic nonstationarity case as the random nonstationarity case was previously studied in the literature (see, e.g., [18]). We use the ratio of the minimum achievable steady-state MSE of each of the algorithms to that of the LMS algorithm as a performance measure.

For the LMF algorithm, this ratio is given, from Table III, by

$$\frac{\zeta_{\min}^{\mathsf{LMS}}}{\zeta_{\min}^{\mathsf{LMF}}} = 3\sigma_v^2\left(\frac{3\sigma_v^6}{(\xi_v^6)^2}\right)^{1/3}. \tag{43}$$

Here, we can see that the ratio depends only on the statistical properties of the measurement noise $v(i)$. For the case of the LMMN algorithm, the same ratio is given by

$$\frac{\zeta_{\min}^{\mathsf{LMS}}}{\zeta_{\min}^{\mathsf{LMMN}}} = \gamma\left(\frac{\gamma\sigma_v^4}{a^2}\right)^{1/3} \tag{44}$$

which is also dependent on the statistical properties of the noise, as well as on the norm mixing parameter δ. We specialize these results for the following noise distributions.

Gaussian Noise: In this case, $\xi_4^v = 3\sigma_v^4$ and $\xi_6^v = 15\sigma_v^6$. Then, we can verify from (43) that

$$\frac{\zeta_{\min}^{\mathsf{LMS}}}{\zeta_{\min}^{\mathsf{LMF}}} = \left(\frac{81}{225}\right)^{1/3} \approx -1.5\text{ dB}.$$

This indicates that the minimum achievable value of steady-state MSE of the LMS algorithm is less than that of the LMF algorithm by approximately 1.5 dB for all values of the noise variance σ_v^2. For the complex case, this value drops to approximately 0.65 dB. For the case of the LMMN algorithm, (44) yields

$$\frac{\zeta_{\min}^{\mathsf{LMS}}}{\zeta_{\min}^{\mathsf{LMMN}}} = \gamma\left(\frac{\sigma_v^4\gamma}{\left(\delta^2\sigma_v^2 + 6\delta\bar{\delta}\sigma_v^4 + 15\bar{\delta}^2\sigma_v^6\right)^2}\right)^{1/3}. \tag{45}$$

Fig. 1 shows a plot of this ratio versus the design parameter δ for various values of σ_v^2. The figure shows that this ratio is always less than unity for all values of δ and σ_v^2. These results reflect the superiority of the LMS algorithm over both the LMF and LMMN algorithms for tracking nonstationary systems in Gaussian noise environments.

Uniform Noise: For a uniformly distributed noise in the interval $[-\Delta, \Delta]$, we have $\sigma_v^2 = \Delta^2/3$, $\xi_4^v = \Delta^4/5$, and $\xi_6^v = \Delta^6/7$. Then, we can verify from (43) that

$$\frac{\zeta_{\min}^{\mathsf{LMS}}}{\zeta_{\min}^{\mathsf{LMF}}} = \left(\frac{49}{9}\right)^{1/3} \approx 2.5\text{ dB}.$$

This indicates that the minimum achievable value of steady-state MSE of the LMF algorithm is less than that of the LMS algorithm by approximately 2.5 dB for uniformly distributed noise. For the complex case, this value drops to approximately 0.1 dB.

TABLE IV
EXPRESSIONS FOR a_1, a_3, AND ROUGH ESTIMATES FOR μ_o AND $\zeta_{\min}$ FOR CARRIER OFFSETS ONLY

ALGORITHM	a_1	a_3	μ_o	ζ_{min}
LMS	$M\sigma_u^2\sigma_v^2$	$4\Omega^2\Vert\mathbf{w}^o\Vert^2/\sigma_u^2$	$\left(\frac{4\,\Omega^2\,\Vert\mathbf{w}^o\Vert^2}{M\sigma_v^2\sigma_u^4}\right)^{1/3}$	$\frac{3}{4}\left(4M^2\sigma_v^4\sigma_u^2\Omega^2\Vert\mathbf{w}^o\Vert^2\right)^{1/3}$
LMMN	$M\sigma_u^2 a$	$\frac{4\Omega^2\Vert\mathbf{w}^o\Vert^2}{\gamma\sigma_u^2}$	$\left(\frac{4\,\Omega^2\,\Vert\mathbf{w}^o\Vert^2}{a\gamma M\sigma_u^2}\right)^{1/3}$	$\frac{3}{4\gamma}\left(\frac{4a^2M^2\sigma_u^2\Omega^2\Vert\mathbf{w}^o\Vert^2}{\gamma}\right)^{1/3}$
LMF	$M\sigma_u^2\xi_v^6$	$\frac{4\Omega^2\Vert\mathbf{w}^o\Vert^2}{3\sigma_v^2\sigma_u^2}$	$\left(\frac{4\,\Omega^2\,\Vert\mathbf{w}^o\Vert^2}{3M\sigma_v^2\xi_v^6\sigma_u^2}\right)^{1/3}$	$\frac{3}{12\sigma_v^2}\left(\frac{4(\xi_v^6)^2M^2\sigma_u^2\Omega^2\Vert\mathbf{w}^o\Vert^2}{3\sigma_v^2}\right)^{1/3}$
SA	$M\sigma_u^2$	$\frac{4\Omega^2\Vert\mathbf{w}^o\Vert^2}{\eta\sigma_u^2}$	$\left(\frac{4\,\Omega^2\,\Vert\mathbf{w}^o\Vert^2}{\eta M\sigma_u^4}\right)^{1/3}$	$\frac{3}{4\eta}\left(\frac{4M^2\sigma_u^2\Omega^2\Vert\mathbf{w}^o\Vert^2}{\eta}\right)^{1/3}$

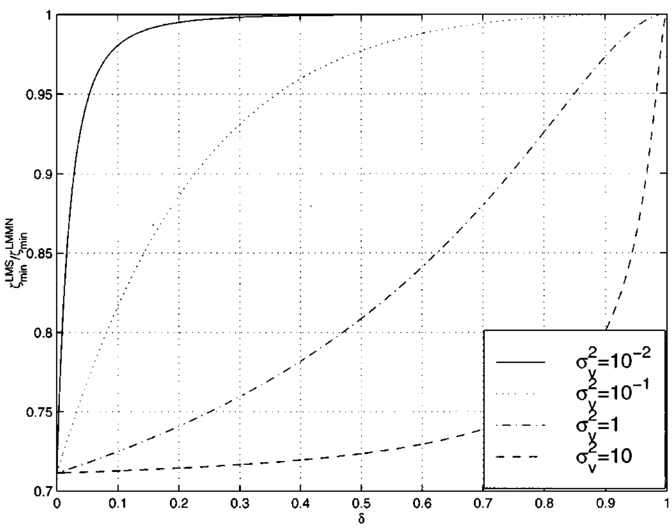

Fig. 1. Comparison of the tracking performance of LMS and LMMN for Gaussian noise.

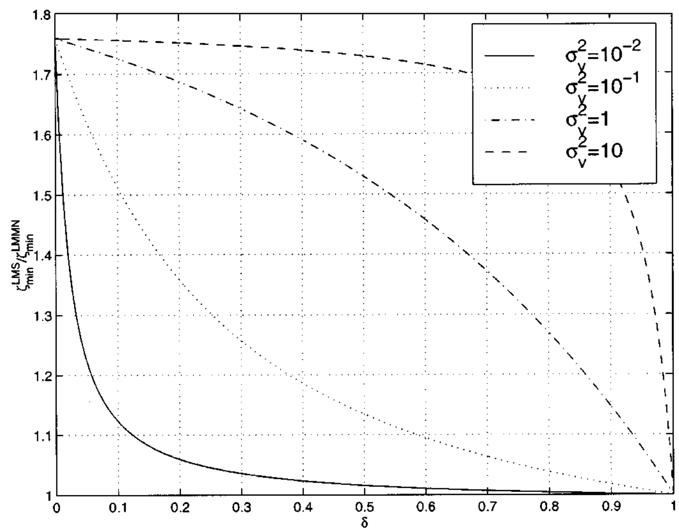

Fig. 2. Comparison of the tracking performance of LMS and LMMN for uniform noise.

Fig. 2 shows a plot of the ratio of the minimum achievable EMSE of the LMS and LMMN algorithms versus the design parameter δ for various values of σ_v^2. The figure shows that this ratio is always larger than unity for all values of δ and σ_v^2. We can also see that $\delta = 0$ results in the best tracking performance, which reflects the superiority of the LMF algorithm in this case.

Mixed Gaussian and Uniform Noise: We now consider the case where the noise is a mix of Gaussian and uniform distributions (for example, a mix of Gaussian system noise and uniformly distributed roundoff errors).[8] Fig. 3 shows the ratio of the minimum achievable EMSE of the LMS and LMMN algorithms versus δ for different values of the system noise variance σ_v^2, which is a combination of Gaussian and uniformly distributed noise with variance ratio 1:3. We can see that in this case, the LMMN algorithm will have the best tracking performance.

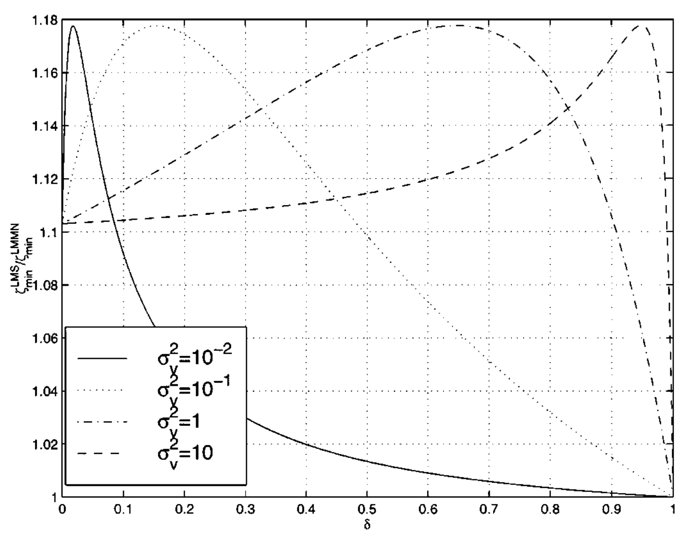

Fig. 3. Comparison of the tracking performance of LMS and LMMN for a mixed Gaussian/uniform noise distribution.

B. Sign Algorithm

We now compare the ability of the sign algorithm to track variations in nonstationary environments with that of the LMS algorithm. We focus only on the cyclic nonstationarity case since the random nonstationarity case was previously studied in the literature (see, e.g., [10]), where it was shown that the LMS is superior to the SA by approximately 1 dB. The ratio of the minimum achievable steady-state EMSE is given, from Table III, by

$$\frac{\zeta_{\min}^{\mathsf{LMS}}}{\zeta_{\min}^{\mathsf{SA}}} = \eta\left(\eta\sigma_v^4\right)^{1/3} = \left(\frac{2}{\pi}\right)^{5/6} \approx -0.16\text{ dB}.$$

This indicates that the minimum achievable value of steady-state EMSE of the LMS algorithm is less than that of the SA algorithm

[8]In communication systems, the noise is usually Gaussian. However, when adaptive algorithms are implemented in finite precision, quantization errors (which are often uniformly distributed) are also added to the system noise, resulting in a mixed noise distribution.

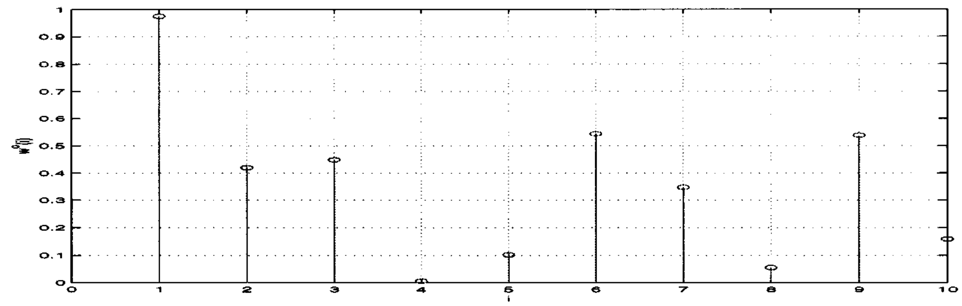

Fig. 4. Unknown system impulse response.

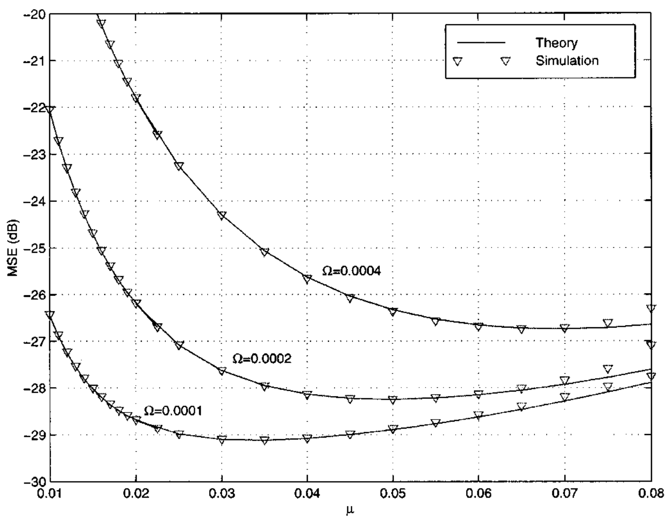

Fig. 5. Theoretical and experimental MSE of the LMS algorithm versus μ for various values of Ω.

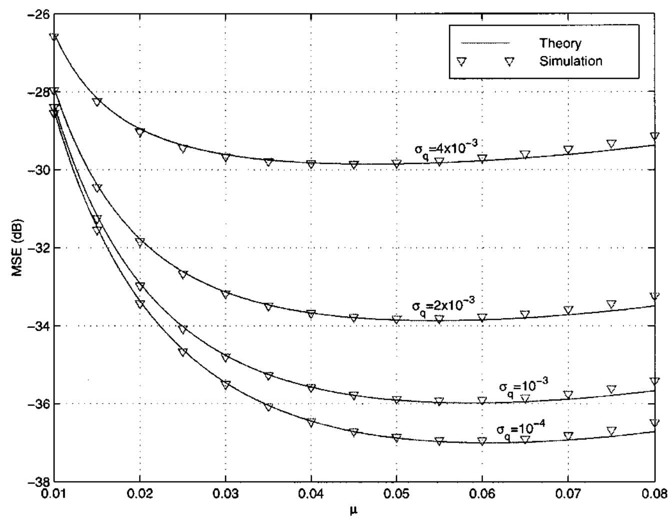

Fig. 6. Theoretical and experimental MSE of the LMS algorithm versus μ for various values of σ_q.

by approximately 0.16 dB for all values of the noise variance σ_v^2. This shows that the ability of the sign algorithm to track random system nonstationarities is very close to that of the LMS algorithm.

V. Simulation Results

A. LMS

Fig. 5 compares the theoretical and experimental MSE of the LMS algorithm for a wide range of step sizes μ and for three different values of the carrier offset Ω (0.0001, 0.0002, 0.0004). In the simulations, we used a random binary phase shift keying (BPSK) input signal of unity variance, a ten-tap unknown system of impulse response shown in Fig. 4[9], $\sigma_v = 3 \times 10^{-2}$, $\alpha = 0.9$, $\sigma_q = 10^{-4}$. Each simulation point is the average of 100 runs with 3000 iterations in each run.

It is clear from Fig. 5 that the theoretical results are a very good match with the simulation results. For $\Omega = 0.0001$, we can see that the experimental MSE possesses a well-defined minimum at $\mu = 0.035$, which is a very good match with the solution of (40) ($\mu = 0.0351$) and close to the estimate provided by (41) ($\mu_o = 0.0381$). We can also see that the minimum achievable MSE is degraded by 0.9 and 2.39 dB, respectively, when Ω is doubled and quadrupled. This reflects that the tracking performance of the algorithm can significantly be affected by the frequency offset Ω, even for very small values of Ω!

Fig. 6 shows the theoretical and experimental EMSE versus μ for $\Omega = 0.001$, $\sigma_v = 10^{-2}$, $\alpha = 0.9$, and various values of σ_q. The figure shows that cyclic nonstationarities are dominant for small values of μ. For example, for $\mu = 0.01$, the EMSE varies within less than 1 dB when σ_q is varied from 10^{-4} to 4×10^{-3}. However, it varies by more than 7 dB for larger values of μ. This can be explained as follows. For small values of μ, cyclic nonstationarities are dominant as the cyclic nonstationarity term in the EMSE is inversely proportional to μ^2, whereas the random nonstationarity term is inversely proportional to μ. Thus, the effect of random nonstationarities is more significant for relatively larger values of μ. Note also that Figs. 5 and 6 show that the effect of carrier offset nonstationarities can be more damaging to the tracking performance of the LMS algorithm as it is inversely proportional to μ^2 and increases with the *square* of the carries offset Ω. On the other hand, the effects of random channel nonstationarities decrease with μ and increase *linearly* with random nonstationarity power $\mathrm{Tr}(\mathbf{Q})$.

Fig. 7(a) and (b) show the real and imaginary parts of the first adaptive filter weight and the real and imaginary parts of the corresponding system weight to be tracked $\mathbf{w}_i^o e^{j\Omega i}$ versus the time index i. It is clear that the filter weight tracks the system weight. Fig. 7(c) and (d) show the real and imaginary parts of

[9]In this figure, the given impulse response represents the constant nonfading portion of the channel.

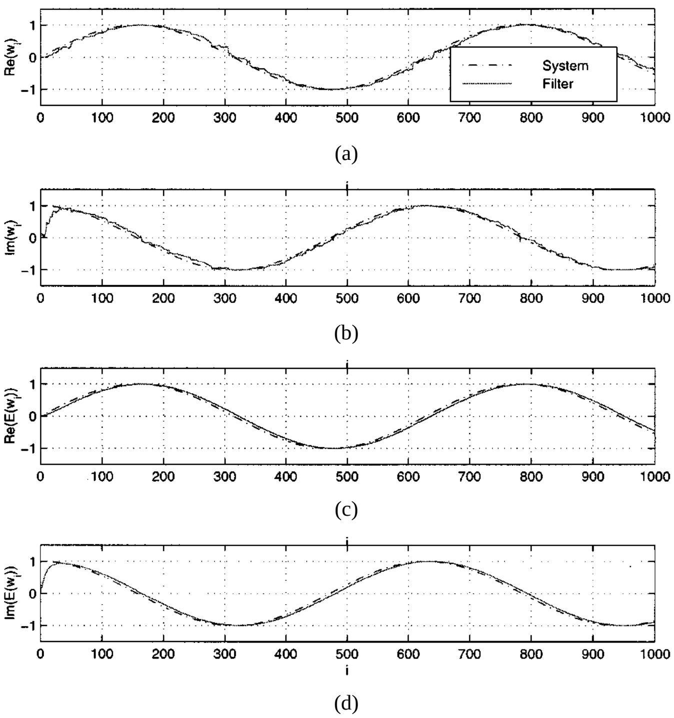

Fig. 7. Adaptive filter and system weight variations with time.

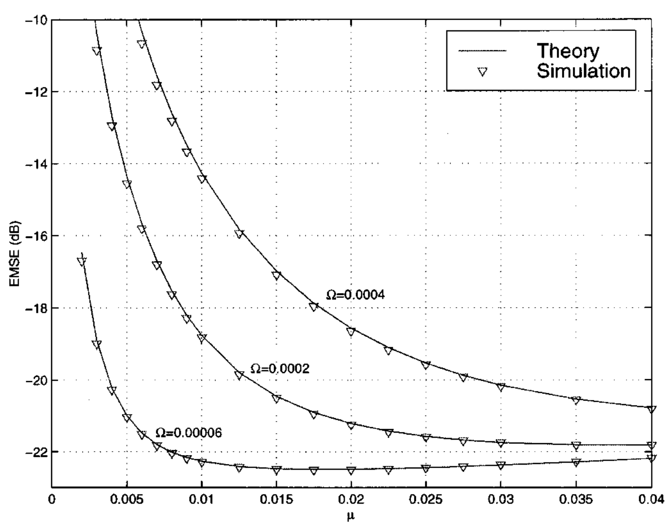

Fig. 8. Experimental and theoretical MSE versus μ for the LMMN algorithm.

the first adaptive filter weight averaged over 100 runs and the real and imaginary parts of the corresponding system weight. This verifies that in steady state, the adaptive filter weight vector converges to the generic form $E(\tilde{\mathbf{w}}_i) = \mathbf{v}e^{j\Omega i}$, which was predicted by the analysis. In this simulation, $\Omega = 0.01$, and $\mu = 0.1$.

B. LMMN

Fig. 8 compares the simulation and theoretical MSE for a wide range of step sizes μ and for three different values of the carrier offset Ω (0.000 06, 0.0002, 0.0004). In the simulations, we used a white Gaussian input signal of unity variance,

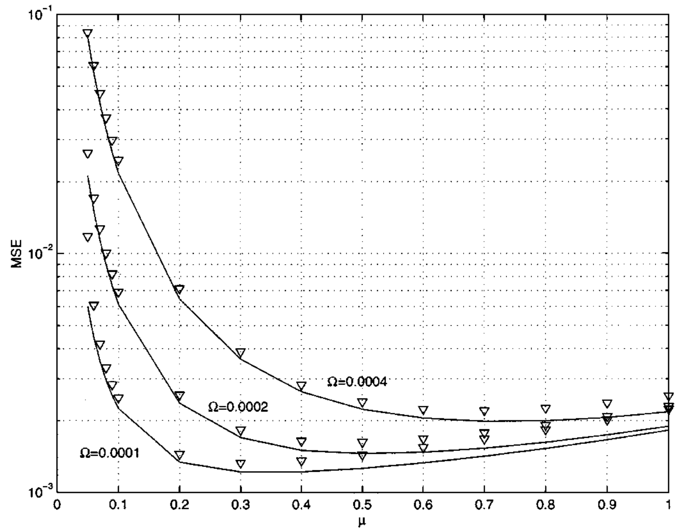

Fig. 9. Theoretical and experimental MSE of the NLMS algorithm versus μ for various values of Ω.

a ten-tap unknown system of impulse response shown in Fig. 4, $\sigma_v = 5 \times 10^{-3}$, and $\mathrm{Tr}(\mathbf{Q}) = 10^{-7}$. Each simulation point is the average of 50 runs with 3000 iterations in each run.

It is clear from Fig. 8 that the theoretical results are a very good match with the simulation results. For $\Omega = 0.000\,06$, we can see that the experimental MSE possesses a well-defined minimum at $\mu = 0.0175$, which is a good match with the expression from Table IV—$\mu_o = 0.0164$. We can also see that the minimum achievable MSE is degraded by approximately 1 dB when Ω is increased to 0.0002 and then 0.0004. This reflects that the tracking performance of the LMMN algorithm is also significantly affected by the frequency offset Ω, even for very small values of Ω!

C. NLMS

Fig. 9 compares the theoretical and experimental MSE of the NLMS algorithm for a wide range of the step-size μ, three different values of the carrier offset Ω (0.0001, 0.0002, 0.0004), and the same simulation conditions of Fig. 5. We can see that the theoretical results are a very good match with the simulation results.

VI. Conclusions

In this paper, we studied the tracking performance of adaptive filters in the presence of two sources of nonstationarities: carrier frequency offsets and random variations. In particular, we derived expressions for the excess-mean-square error that show how the performance is degraded by carrier offsets. We also indicated parameter selection (for step-sizes) to achieve optimal performance.

We may add that the approach can be extended to other scenarios as well, such as the study of the tracking performance of adaptive schemes in finite-precision implementations and the study of adaptive filters of RLS and Gauss-Newton type as well as fractionally spaced blind adaptive schemes (see, e.g., [23], [31]–[33]).

APPENDIX A

Expanding (17) and applying the expectation operator to both of its sides, we obtain

$$E\|\tilde{\mathbf{w}}_{i+1}\|^2 + E\bar{\mu}(i)|e_a(i)|^2 = E\|\tilde{\mathbf{w}}_i\|^2 + E\bar{\mu}(i)|e_p(i)|^2 + 2\text{Re } E\mathbf{c}_i^*\tilde{\mathbf{w}}_{i+1}e^{-j\Omega i} - E\|\mathbf{c}_i\|^2. \quad (46)$$

In steady state (i.e., as $i \to \infty$), we can assume that

$$E\|\tilde{\mathbf{w}}_{i+1}\|^2 = E\|\tilde{\mathbf{w}}_i\|^2. \quad (47)$$

This assumption is equivalent to assuming that the mean square deviation (MSD) converges to a steady-state value. This is a justifiable assumption since our aim is to study the performance of gradient based algorithms after steady state is reached. By imposing (47), we will be able to evaluate the minimum value that we can expect for the MSE at steady state.[10] By imposing (47), we get

$$E\left(\bar{\mu}(i)|e_a(i)|^2\right) = E\left(\bar{\mu}(i)|e_p(i)|^2\right) + 2\text{Re } E\left(\mathbf{c}_i^*\tilde{\mathbf{w}}_{i+1}e^{-j\Omega i}\right) - E\left(\|\mathbf{c}_i\|^2\right).$$

Squaring (12), applying the expectation operator to its sides, and neglecting zero-mean terms, we obtain

$$E\left(\|\mathbf{c}_i\|^2\right) = \text{Tr}(\mathbf{Q}) + |1-e^{j\Omega}|^2\text{Tr}(\mathbf{W}^o) + |1-\alpha e^{j\Omega}|^2\text{Tr}(\boldsymbol{\Theta}).$$

Using (11) and (12), we get

$$\begin{aligned} E\left(\bar{\mu}(i)|e_a(i)|^2\right) =& E\left(\bar{\mu}(i)|e_p(i)|^2\right) + \text{Tr}(\mathbf{Q}) \\ &+ |1-e^{j\Omega}|^2\text{Tr}(\mathbf{W}^o) \\ &+ |1-\alpha e^{j\Omega}|^2\text{Tr}(\boldsymbol{\Theta}) \\ &- 2\text{Re}\Big[\left(1-e^{-j\Omega}\right)\mathbf{w}^{o*} \\ &\qquad \times E\left(\left(\tilde{\mathbf{w}}_i - \mu\mathbf{u}_i^* f_e(i)\right)e^{-j\Omega i}\right)\Big] \\ &- 2\text{Re}\Big[\left(1-\alpha^* e^{-j\Omega}\right) \\ &\qquad \times E\left(\boldsymbol{\theta}_i^*\left(\tilde{\mathbf{w}}_i - \mu\mathbf{u}_i^* f_e(i)\right)e^{-j\Omega i}\right)\Big]. \end{aligned}$$

Substituting (13) into the above expression, we arrive at

$$\begin{aligned} E\left(\bar{\mu}(i)|e_a(i)|^2\right) =& E\bar{\mu}(i)\left|e_a(i) - \frac{\mu}{\bar{\mu}(i)}f_e(i)\right|^2 + \text{Tr}(\mathbf{Q}) \\ &+ |1-e^{j\Omega}|^2\text{Tr}(\mathbf{W}^o) \\ &+ |1-\alpha e^{j\Omega}|^2\text{Tr}(\boldsymbol{\Theta}) \\ &- 2\text{Re}\Big[\left(1-e^{-j\Omega}\right)\mathbf{w}^{o*} \\ &\qquad \times E\left(\left(\tilde{\mathbf{w}}_i - \mu\mathbf{u}_i^* f_e(i)\right)e^{-j\Omega i}\right)\Big] \\ &- 2\text{Re}\Big[\left(1-\alpha^* e^{-j\Omega}\right) \\ &\qquad \times E\left(\boldsymbol{\theta}_i{}^*\left(\tilde{\mathbf{w}}_i - \mu\mathbf{u}_i^* f_e(i)\right)e^{-j\Omega i}\right)\Big]. \end{aligned}$$

[10]We may mention that by averaging analysis, and under some conditions, one can guarantee that there exists a small enough μ for which the filter reaches steady state (see, e.g., [4], [24]); we do not expand on the stability issue here since the objective of this paper is to evaluate filter performance once steady state is reached; see [32], [33].

Expanding the first term on the right-hand side and rearranging terms, (18) will follow.

APPENDIX B

A. LMF and LMMN Algorithms

Here, we need only study the tracking performance of the LMMN algorithm and then obtain the LMF algorithm tracking results as a special case by setting $\delta = 0$ [28]. Introduce, for compactness of notation

$$\begin{aligned} \xi_v^4 =& E|v(i)|^4, \quad \xi_v^6 = E|v(i)|^6 \\ \bar{\delta} =& 1-\delta, \quad \gamma = \delta + 3\bar{\delta}\sigma_v^2. \end{aligned}$$

Now, in steady state and when μ is small enough, it is reasonable to assume that $|e_a(i)|^2 \ll |v(i)|^2$. Applying the expectation operator to both sides of (11) and using (12), A.1 and A.2, it can be shown that if the following condition holds:

$$\mu < \frac{2}{\lambda_{\max}\gamma} \quad (48)$$

expressions (22) and (29) hold, with $\boldsymbol{\Gamma} = \gamma\mathbf{R}$. Substituting (22) and (29) into (18) and using A.1, we obtain

$$\begin{aligned} 2\mu\gamma\zeta^{\text{LMMN}} =& \mu^2 a\text{Tr}(\mathbf{R}) + \mu^2 bE\left(\|\mathbf{u}_i\|^2|e_a(i)|^2\right) \\ &+ |1-e^{j\Omega}|^2\text{ReTr}\left[\mathbf{W}^o\left(\mathbf{I}-2\mathbf{X}\right)\right] \\ &+ |1-\alpha e^{j\Omega}|^2\text{ReTr}\left[\boldsymbol{\Theta}\left(\mathbf{I}-2\alpha^*\mathbf{X}_\alpha\right)\right] \\ &+ \text{Re Tr}\left[\mathbf{Q}\left(\mathbf{I}-2\left(\alpha^*-e^{-j\Omega}\right)\mathbf{X}_\alpha\right)\right] \quad (49) \end{aligned}$$

where

$$\begin{aligned} a &= \left(\delta^2\sigma_v^2 + 2\delta\bar{\delta}\xi_v^4 + \bar{\delta}^2\xi_v^6\right) \\ b &= \left(\delta^2 + 12\delta\bar{\delta}\sigma_v^2 + 15\bar{\delta}^2\xi_v^4\right). \end{aligned}$$

Solving for ζ^{LMMN}, we obtain the expressions given in Table III for the LMMN and LMF.

Here, we may add that for the case of complex-valued data, we replace e^3 by $e|e|^2$ and assume the noise is circular, i.e., $Ev^2(i) = 0$. Then, repeating the above arguments, we find that the EMSE expressions are still valid but with γ and b replaced by

$$\gamma' = \delta + 2\bar{\delta}\sigma_v^2, \quad b' = \delta^2 + 8\delta\bar{\delta}\sigma_v^2 + 9\bar{\delta}\xi_4^v.$$

B. Sign Algorithm

Using the Price theorem[11] [30], we can show that if the following condition holds:

$$\mu < \frac{2}{\lambda_{\max}\eta} \quad (50)$$

where

$$\eta = \sqrt{\frac{2}{\pi(\sigma_v^2 + \zeta)}}$$

[11]For two jointly Gaussian real-valued random variables x and y, we have $E[x \text{ sign}(y)] = \sqrt{2/\pi}\cdot(1/\sigma_y)E[xy]$.

expressions (22) and (29) hold with $\mathbf{\Gamma} = \eta\mathbf{R}$. Substituting (22) and (29) into (18) and using $|e_a(i)|^2 \ll \sigma_v^2$ in steady state, the Price theorem [30], and A.1, we obtain

$$\begin{aligned} 2\mu\eta\zeta^{\mathsf{SA}} =& \mu^2\mathrm{Tr}(\mathbf{R}) + |1-e^{j\Omega}|^2 \mathrm{Re\ Tr}\left[\mathbf{W}^o\left(\mathbf{I}-2\mathbf{X}\right)\right] \\ &+ |1-\alpha e^{j\Omega}|^2 \mathrm{Re\ Tr}\left[\mathbf{\Theta}\left(\mathbf{I}-2\alpha^*\mathbf{X}_\alpha\right)\right] \\ &+ \mathrm{Re\ Tr}\left[\mathbf{Q}\left(\mathbf{I}-2\left(\alpha^*-e^{-j\Omega}\right)\mathbf{X}_\alpha\right)\right]. \end{aligned} \tag{51}$$

Solving for the EMSE, we get the expression given in Table III for the SA.

C. NLMS *Algorithm*

If the following condition holds:

$$\mu < \frac{2}{\lambda_{\max}(\mathbf{R}_N)} \tag{52}$$

where

$$\mathbf{R}_N \triangleq E\left(\frac{\mathbf{u}_i^*\mathbf{u}_i}{\|\mathbf{u}_i\|^2}\right)$$

and $\lambda_{\max}(\mathbf{R}_N)$ is the maximum eigenvalue of $\mathbf{R}_N$, it can be shown that (22) and (29) hold for the NLMS with $\mathbf{\Gamma} = \mathbf{R}_N$. Substituting (22) and (29) into (18) and using A.1–A.3, it can be verified that

$$\begin{aligned} 2\mu E\left(\frac{1}{\|\mathbf{u}_i\|^2}\right)\zeta^{\mathsf{LMS}} =& \mu^2\sigma_v^2 E\left(\frac{1}{\|\mathbf{u}_i\|^2}\right) \\ &+ \mu^2 E\left(\frac{1}{\|\mathbf{u}_i\|^2}\right)\zeta^{\mathsf{NLMS}} \\ &+ |1-\alpha e^{j\Omega}|^2 \mathrm{Re\ Tr}\left[\mathbf{\Theta}\left(\mathbf{I}-2\alpha^*\mathbf{X}_\alpha\right)\right] \\ &+ \mathrm{Re\ Tr}\left[\mathbf{Q}\left(\mathbf{I}-2\left(\alpha^*-e^{-j\Omega}\right)\mathbf{X}_\alpha\right)\right]. \end{aligned} \tag{53}$$

Solving for the EMSE, we get the expression given in Table III for the NLMS.

ACKNOWLEDGMENT

The authors would like to thank Dr. M. Rupp for bringing to their attention the work [11].

REFERENCES

[1] B. Widrow and S. D. Stearns, *Adaptive Signal Processing*. Englewood Cliffs, NJ: Prentice-Hall, 1985.
[2] S. Haykin, *Adaptive Filter Theory*, 3rd ed. Englewood Cliffs, NJ: Prentice-Hall, 1996.
[3] O. Macchi, *Adaptive Processing: The LMS Approach with Applications in Transmission*. New York: Wiley, 1995.
[4] V. Solo and X. Kong, *Adaptive Signal Processing Algorithms*. Englewood Cliffs, NJ: Prentice-Hall, 1995.
[5] B. Widrow, J. McCool, M. G. Larimore, and C. R. Johnson, "Stationary and nonstationary learning characteristics of the LMS adaptive filter," *Proc. IEEE*, vol. 46, pp. 1151–1162, Aug. 1976.
[6] E. Eweda and O. Macchi, "Tracking error bounds of adaptive nonstationary filtering," *Automatica*, vol. 21, no. 3, pp. 293–302, May 1985.
[7] E. Eleftheriou and D. D. Falconer, "Tracking properties and steady-state performance of RLS adaptive filter algorithms," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-34, pp. 1097–1110, Oct. 1986.
[8] S. Marcos and O. Macchi, "Tracking capability of the least mean square algorithm: Application to an asynchronous echo canceler," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-35, pp. 1570–1578, Nov. 1987.
[9] M. Hajivandi and W. A. Gardner, "Measures of tracking performance for the LMS algorithm," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 38, pp. 1953–1958, Nov. 1990.
[10] E. Eweda, "Comparison of RLS, LMS, and sign algorithms for tracking randomly time-varying channels," *IEEE Trans. Signal Processing*, vol. 42, pp. 2937–2944, Nov. 1994.
[11] M. Rupp, "LMS tracking behavior under periodically changing systems," in *Proc. Eur. Signal Process. Conf.*, Island of Rhodes, Greece, Sept. 1998.
[12] A. R. S. Bahai and M. Sarraf, "A frequency offset estimation technique for nonstationary channels," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, vol. 5, Munich, Germany, Apr. 1997, pp. 21–24.
[13] T. K. Tsatsanis, G. B. Giannakis, and G. Zhou, "Estimation and equalization of fading channels with random coefficients," *Signal Process.*, vol. 53, no. 2–3, pp. 211–229, Sept. 1996.
[14] C. Kominakis, C. Fragouli, A. H. Sayed, and R. D. Wesel, "Channel estimation and equalization in frequency selective fading," in *Proc. Asilomar Conf. Signals, Syst., Comput.*, vol. 1, Oct. 1999, pp. 1159–1163.
[15] M. Ghosh, "Blind decision feedbach equalization for teresstial television receivers," in *Proc. Asilomar Conf. Signals, Syst., Comput.*, vol. 1, Oct. 1999, pp. 1159–1163.
[16] S. Cho and V. Mathews, "Tracking analysis of sign algorithm in nonstationary environments," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 38, pp. 2046–2057, Dec. 1990.
[17] E. Eweda, "Optimum step size of the sign algorithm for nonstationary adaptive filtering," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 38, pp. 1897–1901, Nov. 1990.
[18] N. R. Yousef and A. H. Sayed, "Tracking analysis of the LMF and LMMN adaptive algorithms," in *Proc. Asilomar Conf. Signals, Syst., Comput.*, vol. 1, Oct. 1999, pp. 786–790.
[19] A. H. Sayed and M. Rupp, "A time-domain feedback analysis of adaptive algorithms via the small gain theorem," *Proc. SPIE*, vol. 2563, pp. 458–469, Jul. 1995.
[20] M. Rupp and A. H. Sayed, "A time-domain feedback analysis of filtered-error adaptive gradient algorithms," *IEEE Trans. Signal Processing*, vol. 44, pp. 1428–1439, June 1996.
[21] A. H. Sayed and M. Rupp, "Robustness issues in adaptive filtering," in *DSP Handbook*. Boca Raton, FL: CRC, 1998, ch. 20.
[22] N. R. Yousef and A. H. Sayed, "A unified approach to the steady-state and tracking analyzes of adaptive filters," *IEEE Trans. Signal Processing*, vol. 49, pp. 314–324, Feb. 2001.
[23] J. Mai and A. H. Sayed, "A feedback approach to the steady-state performance of fractionally-spaced blind adaptive equalizers," *IEEE Trans. Signal Processing*, vol. 48, pp. 80–91, Jan. 2000.
[24] A. Benveniste, M. Metivier, and P. Priouret, *Adaptive Algorithms and Stochastic Approximations*. New York: Springer-Verlag, 1990.
[25] E. Pfann and R. Stewart, "LMS adaptive filtering with $\Sigma\Delta$ modulated input signals," *IEEE Signal Processing Lett.*, vol. 5, pp. 95–97, Apr. 1998.
[26] J. A. Chambers, O. Tanrikulu, and A. G. Constantindes, "Least mean mixed-norm adaptive filtering," *Electron. Lett.*, vol. 30, no. 19, pp. 1574–1575, Sept. 1994.
[27] O. Tanrikulu and J. A. Chambers, "Convergence and steady-state properties of the least-mean mixed-norm (LMMN) adaptive algorithm," *Proc. Inst. Elect. Eng. Vision, Image Signal Process.*, vol. 143, no. 3, pp. 137–142, June 1996.
[28] E. Walach and B. Widrow, "The least mean fourth (LMF) adaptive algorithm and its family," *IEEE Trans. Inform. Theory*, vol. IT–30, pp. 275–283, Mar. 1984.
[29] V. Mathews and S. Cho, "Improved convergence analysis of stochastic gradient adaptive filters using the sign algorithm," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-35, pp. 450–454, Apr. 1987.
[30] R. Price, "A useful theorem for nonlinear devices having gaussian inputs," *IRE Trans. Inform. Theory*, vol. IT-4, pp. 69–72, June 1958.
[31] N. R. Yousef and A. H. Sayed, "A feedback analysis of the tracking performance of blind adaptive equalization algorithms," in *Proc. Conf. Decision Contr.*, vol. 1, Phoenix, AZ, Dec. 1999, pp. 174–179.
[32] A. H. Sayed and T. Y. Al-Naffouri, "Mean-square analysis of normalized leaky adaptive filters," in *Proc. ICASSP*, vol. VI, Salt Lake City, UT, May 2001.
[33] T. Y. Al-Naffouri and A. H. Sayed, "Transient analysis of adaptive filters," in *Proc. ICASSP*, vol. VI, Salt Lake City, UT, May 2001.

Nabil R. Yousef (S'98) received the B.S. and M.S. degrees in electrical engineering from Ain Shams University, Cairo, Egypt, in 1994 and 1997, respectively, and the Ph.D. in electrical engineering from the University of California, Los Angeles, in 2001.

He is currently a Senior Research Staff at Broadcom Corp., Irvine, CA. His research interests include adaptive filtering, equalization, CDMA systems, and wireless location.

Dr. Yousef is recipient of a 1999 Best Student Paper Award at an international meeting for work on adaptive filtering and of a 1999 NOKIA Fellowship Award.

Ali H. Sayed (F'01) received the Ph.D. degree in electrical engineering in 1992 from Stanford University, Stanford, CA.

He is Professor of electrical engineering at the University of California, Los Angeles. He has over 170 journal and conference publications, is coauthor of the research monograph *Indefinite Quadratic Estimation and Control* (Philadelphia, PA: SIAM, 1999) and of the graduate-level textbook *Linear Estimation* (Englewood Cliffs, NJ: Prentice-Hall, 2000). He is also coeditor of the volume *Fast Reliable Algorithms for Matrices with Structure* (Philadelphia, PA: SIAM, 1999). He is a member of the editorial boards of the *SIAM Journal on Matrix Analysis and Its Applications*, of the *International Journal of Adaptive Control and Signal Processing*, and has served as coeditor of special issues of the journal *Linear Algebra and Its Applications*. He has contributed several articles to engineering and mathematical encyclopedias and handbooks and has served on the program committees of several international meetings. He has also consulted with industry in the areas of adaptive filtering, adaptive equalization, and echo cancellation. His research interests span several areas, including adaptive and statistical signal processing, filtering and estimation theories, equalization techniques for communications, interplays between signal processing and control methodologies, and fast algorithms for large-scale problems. To learn more about his work, visit the website of the UCLA Adaptive Systems Laboratory at http://www.ee.ucla.edu/asl.

Dr. Sayed is a recipient of the 1996 IEEE Donald G. Fink Award. He is Associate Editor of the IEEE TRANSACTIONS ON SIGNAL PROCESSING and is a member of the technical committees on Signal Processing Theory and Methods (SPTM) and on Signal Processing for Communications (SPCOM), both of the IEEE Signal Processing Society.