Measures of Tracking Performance for the LMS Algorithm

MEHRDAD HAJIVANDI AND WILLIAM A. GARDNER, SENIOR MEMBER, IEEE

Abstract—Two measures of tracking performance for the LMS algorithm are compared and contrasted. These are the conventional time-average or temporal mean of the nonstationary mean-squared error (MSE) in excess of the minimum attainable MSE, and the novel temporal root-mean-squared value of the excess MSE, which takes into account the temporal variance as well as the temporal mean of the non-stationary MSE. These measures are evaluated for the LMS algorithm applied to two time-variant system identification problems, one involving a random Markov system and the other a periodic system. Optimal step-size parameters and minimum misadjustments are evaluated. It is shown that the conventional time-average performance measure is adequate only when the degree of nonstationarity is sufficiently low. For higher degrees of nonstationarity, the time-average performance measure can be misleading in studies of the tracking behavior of the LMS algorithm.

I. INTRODUCTION

THE LMS algorithm for adaptive adjustment of an N-vector of filter weights W is given by

$$W(i + 1) = W(i) + \mu e(i) X(i) \qquad (1)$$

where μ is a step-size parameter,[1] $X(i)$ is the filter input vector, $e(i) = d(i) - \hat{d}(i)$ is the error between the desired quantity $d(i)$ and the filter output $\hat{d}(i) = W^T(i) X(i)$, and $e(i) X(i)$ is half the negative gradient of the squared error $e^2(i)$ (with respect to the weight vector $W(i)$).

Previous studies of the tracking performance of the LMS algorithm (and related stochastic-gradient descent algorithms) operating in nonstationary environments have used the time-average of the nonstationary instantaneous mean-squared error produced by the algorithm as the measure of performance [1]–[3]. Typical graphs of nonstationary instantaneous mean-squared error $\epsilon(i) \triangleq E\{e^2(i)\}$ (where $E\{\cdot\}$ denotes probabilistic expectation) for the LMS algorithm are shown in Figs. 1 and 2. That in Fig. 1 is for a randomly varying nonstationarity, and that in Fig. 2 is for a periodically varying nonstationarity (these are described in Section III). It is clear from these graphs that there can be substantial fluctuation in the mean-squared error about its time-averaged value

$$\langle \epsilon(i) \rangle \triangleq \lim_{Z \to \infty} \frac{1}{Z} \sum_{i=0}^{Z-1} \epsilon(i). \qquad (2)$$

[1]The step-size μ differs from that in [1]–[3] by a factor of two.

Fig. 1. Experimental mean-squared error $\epsilon(i)$, as a function of data sample i, for an ensemble size of 300 from model 1, with DNS = 1/512, SNR = 50, $N = 25$, $\mu = 0.018$.
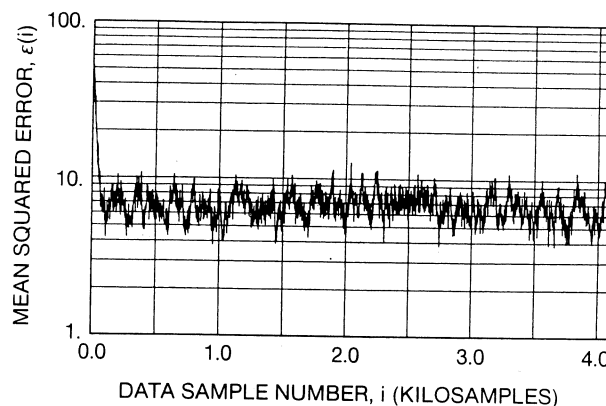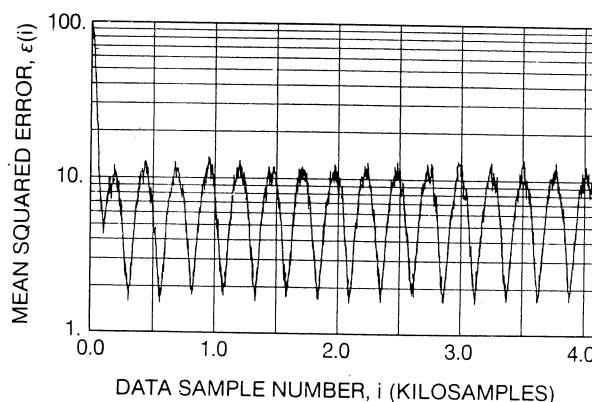

Fig. 2. Experimental mean-squared error $\epsilon(i)$, as a function of data sample i, for an ensemble size of 300 from model 2, with DNS = 1/512, SNR = 50, $N = 25$, $\mu = 0.012$.

This suggests that use of only the time-averaged value for tasks such as step-size optimization for the purpose of predicting optimum-performance might be misleading. Although some investigations of tracking performance have worked directly with the instantaneous mean-squared error (e.g., [5], [6]), the results obtained have understandably been more limited (e.g., restriction to only bounds on performance) than when some type of time average is used. See also [7]. Other studies have used stationary stochastic models for the nonstationarity (e.g., a stochastic process model for the time-variant impulse-response of an unknown system to be identified), and have adopted the steady-state expected value (e.g., over the ensemble of random systems) of the mean-squared error

(e.g., between the outputs of the unknown system and the model, where the expectation used to obtain this mean-squared error is over the ensemble of inputs to the system and its model) as a performance measure [8]. For an ergodic stochastic process model of the nonstationarity, this is equivalent to using the time average of the mean-squared error. An overview of and bibliography on analysis of adaptation and tracking for system identification is given in [9].

These observations suggest using the temporal root-mean-squared value of the instantaneous excess mean-squared error

$$\epsilon_*(i) \triangleq \epsilon(i) - \epsilon_0(i) \tag{3}$$

where $\epsilon_0(i)$ is the minimum attainable value of $\epsilon(i)$, with respect to the weight vector $W(i)$ being adjusted by the LMS algorithm. The temporal root-mean-squared excess mean-squared error

$$\epsilon_{RMS} \triangleq \left\langle \epsilon_*^2(i) \right\rangle^{1/2}$$
$$= \left[\epsilon_{AVE}^2 + \text{var}(\epsilon_*) \right]^{1/2} \tag{4}$$

where

$$\epsilon_{AVE} \triangleq \left\langle \epsilon_*(i) \right\rangle \tag{5}$$

takes into account the temporal variance $\text{var}(\epsilon_*)$ about the temporal mean ϵ_{AVE} of $\epsilon_*(i)$. The objective of this paper is to compare the results of performance studies based on the two alternative performance measures ϵ_{RMS} and ϵ_{AVE}.

The investigation focuses on the same type of time-varying system identification problem considered in previous studies of the nonstationary learning characteristics of the LMS algorithm [1]–[3], and considers the same random Markov model for time-variation used in [1]–[3] as well as a periodic model. It is shown that the discrepancy between optimum step sizes obtained for minimization of ϵ_{AVE} and ϵ_{RMS} can be large for a high degree of nonstationarity, and this can lead to misinterpretations of the behavior of the LMS algorithm. Similar results are obtained for the minimum values of ϵ_{RMS} and ϵ_{AVE}. It is concluded that ϵ_{AVE}, which is considerably more analytically tractable, is adequate only if the degree of nonstationarity is not too high. It can be misleading when the degree of nonstationarity is indeed high, as considered in [1], [6], [8].

II. Formulas for Average and RMS Mean-Squared Error

For the system identification problem of interest, we have

$$d(i) = \bar{W}^T(i) X(i) + n(i) \tag{6}$$

where $\bar{W}(i)$ is the sequence of unknown time-variant weight vectors corresponding to the unit-pulse response of the unknown system, which is assumed to have memory length less than or equal to N, $X(i) = [x(i), x(i -$

$1), x(i - 2), \cdots, x(i - N + 1)]^T$ is the N-vector of samples of the system excitation $x(i)$, which has variance σ_x^2, and $n(i)$ is measurement noise, which has variance σ_n^2. It is assumed that $x(i)$ and $n(i)$ are independent stationary white Gaussian sequences. Two models for the time-variations of the unknown-system weight vector $\bar{W}(i)$ are considered. In model 1, the N elements of the weight vector $\bar{W}(i)$ fluctuate independently of each other, with steady state variance σ_w^2, according to first-order Markov time-series

$$\bar{w}_n(i + 1) = r\bar{w}_n(i) + v_n(i) \tag{7}$$

where $v_n(i)$ are zero-mean i.i.d. sequences. In steady state ($i \rightarrow \infty$), (7) yields the convolution $\bar{w}_n(i) = r^i \otimes v_n(i)$. Thus, as $r \rightarrow 1$, $\bar{w}_n(i)$ becomes very smooth (low degree of nonstationarity), but as $r \rightarrow 0$, $\bar{w}_n(i)$ becomes very erratic (high degree of nonstationarity).

In model 2, the N elements fluctuate jointly according to the periodicity

$$\bar{W}(i) = q(i) W_* \tag{8}$$

where the constant vector W_* is arbitrary, and the periodic factor $q(i)$ is given by the Fourier series

$$q(i) = \sum_{k=0}^{K} a_k \cos(2\pi ik/T) \tag{9}$$

for arbitrary Fourier coefficients a_k (to be specified in the sequel). In general, as $K \rightarrow T$, the bandwidth of $q(i)$ approaches its maximum (high degree of nonstationarity), but as $K \rightarrow 0$, the bandwidth approaches zero (zero degree of nonstationarity).

For the purpose of obtaining a closed form analytical solution for the excess mean-squared error, it is assumed that $X(i)$ and $d(i)$ are each zero-mean independent sequences. However, in view of the definition of $X(i)$, we can see that this is literally impossible. Nevertheless, this significantly simplifying assumption, which is commonly made in analyses of the LMS algorithm, produces theoretical results that typically agree quite closely with simulations when the step size is sufficiently small (cf. [1]). This is corroborated in this paper.

The objective of this section is to present explicit formulas for the average and RMS misadjustments of the LMS algorithm. These misadjustments are defined by $M_{AVE} \triangleq \epsilon_{AVE}/\epsilon_{min}$ and $M_{RMS} \triangleq \epsilon_{RMS}/\epsilon_{min}$, where ϵ_{min} is the average of $\epsilon_0(i)$, $\epsilon_{min} \triangleq \left\langle \epsilon_0(i) \right\rangle$, and $\epsilon_0(i)$ is given by [4]

$$\epsilon_0(i) = \sigma_d^2(i) - P^T(i) R^{-1}(i) P(i) \tag{10}$$

in which $\sigma_d^2(i)$ is the time-variant variance of $d(i)$, $R(i)$ is the time-variant covariance matrix for $X(i)$ (but is simply $R(i) = \sigma_x^2 I$ for the particular system identification problems considered here), and $P(i)$ is the time-variant crosscovariance vector for $X(i)$ and $d(i)$. The weight vector that yields the minimum mean-squared error $\epsilon_0(i)$ is given by [4]

$$W_0(i) = R^{-1}(i) P(i). \tag{11}$$

Since the type of system identification problem considered here is the same as that studied in [1], the general recursion for the excess mean-squared error (3) derived in [1] applies here

$$\epsilon_*(i + 1) = \gamma \epsilon_*(i) + \beta(i). \tag{12}$$

In this recursion

$$\gamma = 1 - 2\mu\sigma_x^2 + \mu^2(N + 2)\sigma_x^4 \tag{13}$$

and $\beta(i) = \beta_\nabla + \beta_\Delta(i)$, where

$$\beta_\nabla = \mu^2 N \sigma_n^2 \sigma_x^4 \tag{14}$$

$$\beta_\Delta(i) = \sigma_x^2 [\Delta^T(i) - 2(1 - \mu\sigma_x^2) \overline{V}^T(i)] \Delta(i). \tag{15}$$

In (15)

$$\overline{V}(i + 1) = (1 - \mu\sigma_x^2)\overline{V}(i) - \Delta(i) \tag{16}$$

where

$$\Delta(i) = W_0(i + 1) - W_0(i). \tag{17}$$

As explained in [1], the component β_∇ in the driving term $\beta(i)$ in the recursion (12) is due to gradient noise, and the component $\beta_\Delta(i)$ is due to nonstationarity. β_∇ vanishes if there is no gradient noise and $\beta_\Delta(i)$ vanishes if there is no nonstationarity, because then $\Delta(i) \equiv 0$.

The average value of the solution $\epsilon_*(i)$ to (12) is easily obtained by simply equating the average values of both sides of (12). This leads to

$$\epsilon_{\text{AVE}} = \frac{\beta_{\text{AVE}}}{1 - \gamma}. \tag{18}$$

The RMS value can be obtained using standard methods for linear time-invariant recursions driven by random stationary time-series [4] or periodic time-series. The result for model 1 is

$$\epsilon_{\text{RMS}}^2 = \sum_{m=-\infty}^{\infty} \hat{R}_\beta(m) \frac{\gamma^{|m|}}{1 - \gamma^2}$$

$$= \int_{-1/2}^{1/2} \hat{S}_\beta(f) |e^{j2\pi f} - \gamma|^{-2} df \tag{19}$$

where $\hat{R}_\beta(m) \triangleq \langle \beta(i + m) \beta(i) \rangle$ is the autocorrelation of the driving sequence $\beta(i)$ and

$$\hat{S}_\beta(f) \triangleq \sum_{m=-\infty}^{\infty} \hat{R}_\beta(m) e^{-j2\pi mf} \tag{20}$$

is the corresponding spectral density. The quantity $[e^{j2\pi f} - \gamma]^{-1}$ is the transfer function for the first-order recursion (12).

The result for model 2 is the same as (19) and in this case reduces to

$$\epsilon_{\text{RMS}}^2 = \sum_{k=0}^{T-1} |e^{j2\pi k/T} - \gamma|^{-2} |b_k|^2 \tag{21}$$

where b_k are the Fourier coefficients of the periodic component β_p of the asymptotically periodic sequence $\beta(i)$

$$b_k \triangleq \frac{1}{T} \sum_{i=0}^{T-1} \beta_p(i) e^{-j2\pi ik/T}. \tag{22}$$

It can be shown using (13)–(18) and $\epsilon_0^2(i) = \sigma_n^2$ that for model 1 we have the average misadjustment (cf. [1])

$$M_{\text{AVE}} = \left[\frac{1}{s - \mu N\sigma_x^2/2} \right] \left[\mu N\sigma_x^2/2 + \frac{\rho(1 - r)}{1 - rs} \right] \tag{23}$$

where $s \triangleq 1 - \mu\sigma_x^2$ and $\rho \triangleq N\sigma_x^2\sigma_w^2/\sigma_n^2$. The parameter ρ is the unknown-system output SNR. It also can be shown using (13)–(20) and $\epsilon_0(i) = \sigma_n^2$, with considerable tedious calculation [10], that for model 1 we have the RMS misadjustment

$$M_{\text{RMS}}^2 = M_{\text{AVE}}^2 + \left[\frac{4\rho^2(1 - r^2)}{(1 - rs)(1 - \gamma^2)} \right]$$

$$\cdot \left[\frac{(1 - s)^2}{1 - rs} + \frac{1 - s + sr + s^2r + 2s^2}{1 + s} \right.$$

$$+ \frac{\gamma(1 - r)^2(1 + rs)(1 + s/r)}{(1 - \gamma r^2)(1 - s/r)}$$

$$\left. + \frac{\gamma(1 - s)(r^2 - 1)(1 - s^3r)s/r}{(1 - \gamma rs)(1 + s)(1 - rs)(1 - s/r)} \right].$$

$$\tag{24}$$

Similarly, it can be shown using (13)–(18) and $\epsilon_0(i) = \sigma_n^2$ that for model 2, with $W_* = [1, 1, 1, \cdots, 1]^T$, we have the average misadjustment [10]

$$M_{\text{AVE}} = \left[\frac{1}{s - \mu N\sigma_x^2/2} \right] \left[\mu N\sigma_x^2/2 + \frac{\rho\eta(0)}{\mu\sigma_x^2(K + 2)T} \right] \tag{25}$$

where $\rho \triangleq \sigma_x^2 N(K + 2)/2\sigma_n^2$ and

$$\eta(k) \triangleq \sum_{i=0}^{T-1} f(i)[f(i) - 2sg(i)]e^{-j2\pi ik/T} \tag{26}$$

in which

$$f(i) \triangleq q(i + 1) - q(i)$$

$$= \frac{1}{2} \sum_{k=0}^{T-1} (a_k + a_{T-k})(e^{j2\pi k/T} - 1)e^{j2\pi ki/T} \tag{27}$$

and

$$g(i) \triangleq \sum_{k=0}^{T-1} \frac{\frac{1}{2}(a_k + a_{T-k})(e^{j2\pi k/T} - 1)}{s - e^{j2\pi k/T}} e^{j2\pi ki/T} \tag{28}$$

where $a_T \triangleq a_0$. The parameters $\eta(k)$ in (26)–(28) are the normalized Fourier coefficients of the periodic part of the asymptotically periodic component $\beta_\Delta(i)$ due to nonstationarity in the driving term $\beta(i)$ in the recursion (12) for the excess MSE.

It can also be shown using (13)–(22) and $\epsilon_0 = \sigma_n^2$ that for model 2 we have the RMS misadjustment [10]

$$M_{\text{RMS}}^2 = M_{\text{AVE}}^2 + \frac{4\rho^2\phi}{T^2(K + 2)^2} \tag{29}$$

where

$$\phi \triangleq \sum_{k=1}^{T-1} \left| \left| \eta(k) \right|^2 \left| e^{j2\pi k/T} - \gamma \right|^{-2}. \tag{30}$$

III. COMPARISON OF AVERAGE AND RMS PERFORMANCES

Formulas (23)–(25) and (29) have been used to evaluate misadjustment performances M_{AVE} and M_{RMS} for various values of the parameters N (the number of weights), ρ (the SNR), μ (the step size), and DNS (the degree of nonstationarity). Since the only effect of σ_x^2 is to scale μ, σ_x^2 was set equal to unity. For model 1, a useful measure of DNS is DNS $= 1 - r$. For model 2, we consider the particular case where $a_k = 1$, $k = 0, 1, 2, \cdots, K$, in which case a useful measure of DNS is DNS$' = K/T$. (See relevant comments in Section II and in [1] for justification of these two definitions of degree of nonstationarity.)

Fig. 3 shows graphs of the step-size values, denoted by μ_{AVE}^0 and μ_{RMS}^0, that minimize the misadjustments M_{AVE} and M_{RMS}, respectively, for model 1. These are shown as functions of DNS. Similar results are shown in Fig. 4 for model 2. It can be seen that for low DNS μ_{AVE}^0 and μ_{RMS}^0 are comparable, but for high DNS μ_{AVE}^0 can be much smaller than μ_{RMS}^0. However, both μ_{AVE}^0 and μ_{RMS}^0 display similar behavior that shows that as DNS increases, the best step size also increases until DNS is so large that the nonstationarity can no longer be tracked well, in which case the best step size decreases with further increases in DNS in order to reduce the effects of gradient noise and better estimate the *time-averaged* nonstationary system. Thus, the effects of tracking lag are considered dominant to the left of the peak in the optimum-step-size curve, whereas the effects of gradient noise are considered dominant to the right. It can be seen that the peaks in the μ_{RMS}^0 curves are always to the right of the corresponding peaks in the μ_{AVE}^0 curves. This reflects the fact that the root-mean-squared MSE is more sensitive to nonstationarity than time-averaged MSE is. Also, the fact that μ_{AVE}^0 goes to zero (in Fig. 3) or becomes undefined (in Fig. 4) when DNS becomes sufficiently large is misleading. The more appropriate optimum step size μ_{RMS}^0, which does not ignore the temporal variance of the nonstationary mean-squared error, gradually decreases as DNS increases but never goes to zero or becomes undefined.

Fig. 5 shows M_{AVE}^0 and M_{RMS}^0, which are the minimum values (with respect to the step size μ) of M_{AVE} and M_{RMS}, as functions of DNS for model 1. Similar results are shown in Fig. 6 for model 2. It can be seen from these graphs that the largest discrepancies (e.g., factors of 2 or 3) between M_{AVE}^0 and M_{RMS}^0 occur for large values of DNS where performance is poorest. For moderate to small values of DNS, the discrepancies are small to moderate.

In order to confirm these theoretical predictions of performance, simulations were performed for several sets of the parameter values DNS, SNR, and N, and the resultant time-average and RMS misadjustments were measured for various values of the step-size parameter μ. In order to
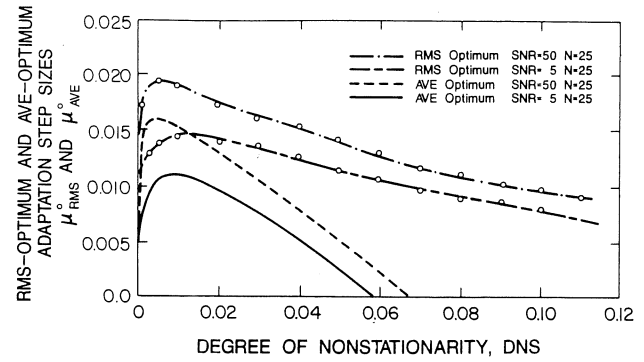

Fig. 3. RMS-optimum and AVE-optimum adaptation step sizes, μ_{AVE}^0 and μ_{RMS}^0, as functions of degree of nonstationarity DNS, for model 1. (Data points from numerical solution shown by symbol o.)
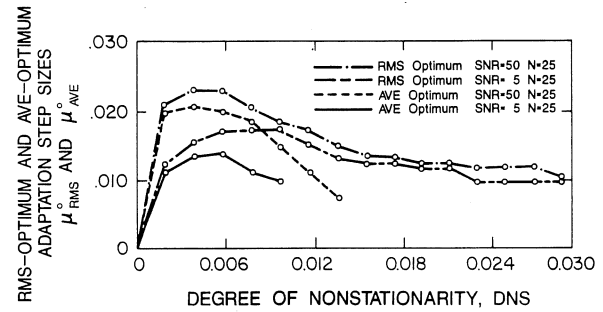

Fig. 4. RMS-optimum and AVE-optimum adaptation step sizes μ_{AVE}^0 and μ_{RMS}^0, as functions of degree of nonstationary DNS, for model 2. (Data points from numerical solution shown by symbol o.)
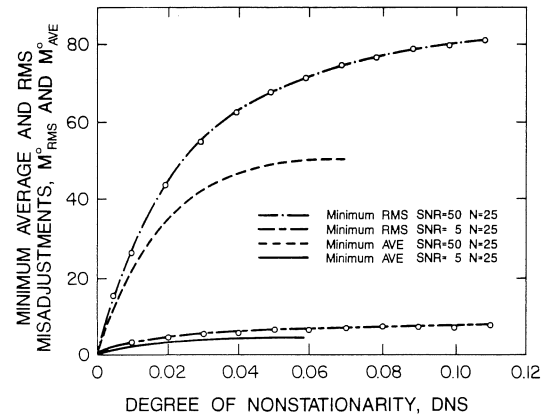

Fig. 5. Minimum average and RMS misadjustments M_{AVE}^0 and M_{RMS}^0, as functions of degree of nonstationarity DNS, for model 1. (Data points from numerical solutions shown by symbol o.)

measure the mean-squared error, an ensemble of 300 random sample paths of $x(i)$ and $n(i)$, both from unity-variance white Gaussian noise generators, was averaged over. In order to measure the time-averaged and RMS values of the excess mean-squared error $\epsilon_*(i)$, the 3583 time-samples from $i = 512$ to $i = 4095$ were used for averaging (initial transients had died away by time $i = 512$). For model 1, each of the N weight sequences in the unknown system weight vector $\tilde{W}(i)$ was an independent sample path of a Gauss–Markov process. These N sample paths were fixed throughout the ensemble of excitation sequences $x(i)$ and measurement noise sequences $n(i)$,
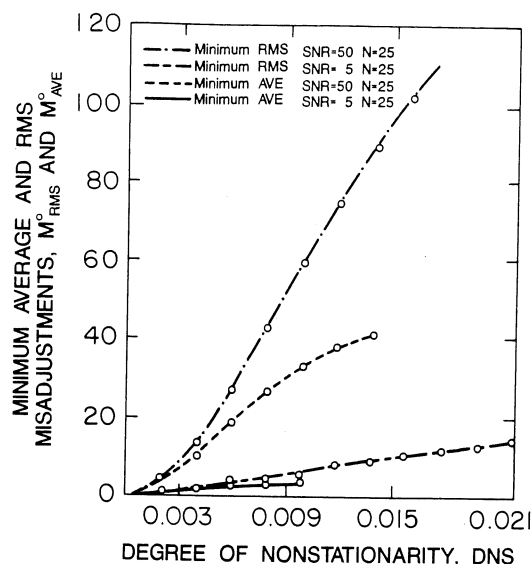
Fig. 6. Minimum average and RMS misadjustments M_{AVE}^0 and M_{RMS}^0, as function of degree of nonstationarity DNS, for model 2. (Data points from numerical solution shown by symbol o.)
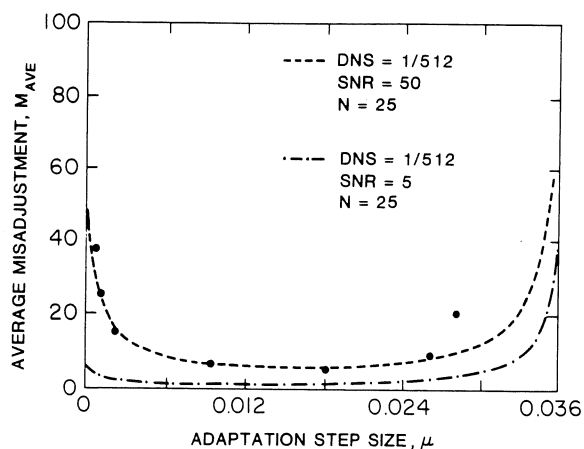

Fig. 7. Average misadjustment M_{AVE}, as a function of adaptation step size μ, for model 1. (Simulation data points for SNR = 50 are superimposed on theoretical curve.)

Fig. 8. RMS misadjustment M_{RMS}, as a function of adaptation step size μ, for model 1. (Simulation data points for SNR = 50 are superimposed on theoretical curve.)

Fig. 9. Average misadjustment M_{AVE}, as a function of adaptation step size μ, for model 2. (Simulation data points for SNR = 50 are superimposed on theoretical curve.)

Fig. 10. RMS misadjustment M_{RMS}, as a function of adaptation step size μ, for model 2. (Simulation data points for SNR = 50 are superimposed on theoretical curve.)

and similarly for model 2 except that for the periodic weight sequences specified by (8), (9), the values $a_k = 1$ $(0 \le k \le K)$ and $W_* = [1, 1, 1, \cdots, 1]^T$ were used. The results are shown in Figs. 7–10. It can be seen that agreement between theory and simulation is quite good. However, the discrepancies that do exist lend further support to the conclusion that the differences between the actual values of μ_{AVE}^0 and μ_{RMS}^0 are not significant for low-to-moderate degree of nonstationarity. All of the performance results presented in Figs. 3–10 are for the parameter values $N = 25$, SNR = 5, 50, and (in Figs. 7–10) DNS = 1/512. Very similar results, which are not presented here, were obtained for the parameter values $N = 5$, SNR = 1, 10, and DNS = 1/256 [10].

IV. Conclusion

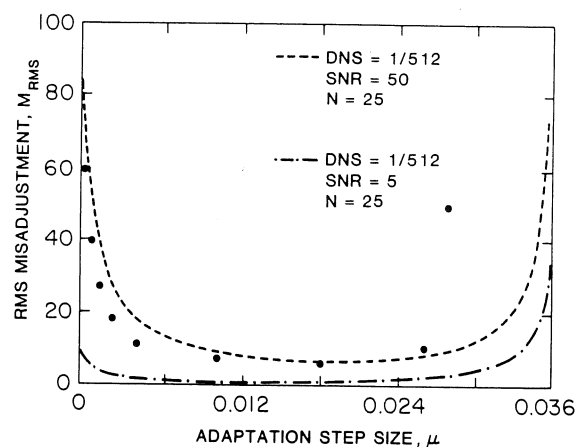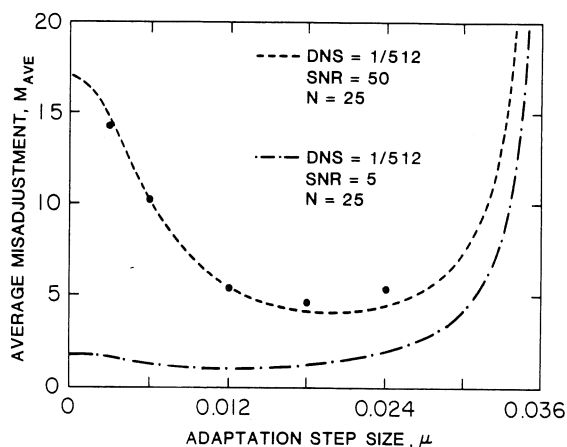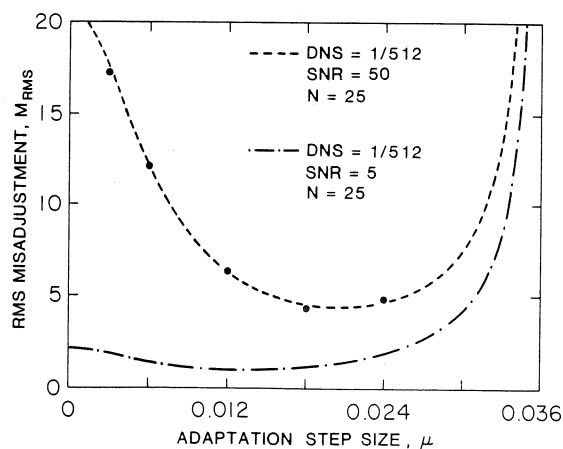In conclusion, the time-averaged MSE, which is considerably more analytically tractable than the temporal root-mean-squared MSE, is an adequate measure of track-ing performance of the LMS algorithm, provided that the degree of nonstationarity is sufficiently low. However, for high degrees of nonstationarity, the temporal root-mean-squared MSE can be as much as two-to-three times larger than the time-averaged MSE. Furthermore, the transition

into the gradient-noise-dominated region of performance (where nonstationarity is simply averaged rather than tracked) is seen to occur much more slowly than the time-averaged-mean-squared-error measure of performance would suggest. This is a result of the fact that the time-averaged measure of performance is less sensitive to non-stationarity, because it ignores the temporal variance of MSE. As a result, the behavior of the step-size optimized LMS algorithm that is predicted by using the time-averaged MSE (as in [1]–[3]) is misleading for high degrees of nonstationarity. In particular, the optimum step size based on the time-averaged MSE goes to zero (or becomes undefined) as the degree of nonstationarity increases suggesting that the best performance will be obtained by minimizing the effects of gradient noise and forfeiting all tracking capability by identifying the time average of the nonstationary system. On the other hand, the behavior of the step-size optimized LMS algorithm that is predicted by using the temporal root-mean-squared MSE is quite different. The optimum step size decreases only gradually as the degree of nonstationarity increases, showing that the effects of tracking error should never be completely ignored.

The inadequacy of the more tractable average misadjustment as a measure of performance for the case of a high degree of nonstationarity is analogous to the well-known inadequacy of the more tractable mean error in the weight-vector for the case of large step size. In both situations, ignoring the error variance can lead to erroneous conclusions about the behavior of the LMS algorithm.

ACKNOWLEDGMENT

REFERENCES

[1] W. A. Gardner, "Nonstationary learning characteristics of the LMS algorithm," *IEEE Trans. Circuits Syst.*, vol. CAS-34, pp. 1199–1207, Oct. 1987.
[2] B. Widrow, J. M. McCool, M. G. Larimore, and C. R. Johnson, Jr., "Stationary and nonstationary learning characteristics of the LMS adaptive filter," *Proc. IEEE*, vol. 64, pp. 1151–1162, Aug. 1976.
[3] B. Widrow and E. Walach, "On the statistical efficiency of the LMS algorithm with nonstationary inputs," *IEEE Trans. Inform. Theory*, vol. IT-30, pp. 211–221, Mar. 1984.
[4] W. A. Gardner, *Introduction to Random Processes with Applications to Signals and Systems*. New York: Macmillan, 1985 (2nd ed., McGraw-Hill, 1990).
[5] D. C. Farden, "Tracking properties of adaptive signal processing algorithms," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-29, pp. 439–446, June 1981.
[6] E. Eweda and O. Macchi, "Tracking error bounds of adaptive nonstationary filtering," *Automatica*, vol. 21, pp. 293–302, 1985.
[7] A. Benveniste, "Design of adaptive algorithms for the tracking of time-varying systems," *Int. J. Adaptive Contr. Signal Processing*, vol. 1, pp. 3–29, 1987.
[8] O. Macchi, "Optimization of adaptive identification for time-varying filters," *IEEE Trans. Automat. Contr.*, vol. AC-31, pp. 283–287, Mar. 1986.
[9] L. Ljung and S. Gunnarsson, "Adaptation and tracking in system identification," Rep. LiTH-ISY-I-0955, Dep. Electrical Engineering, Linköping Univ., S-581 83, Linköping, Sweden, 1988.
[10] M. Hajivandi, "Tracking performance of stochastic gradient algorithms for nonstationary processes," Ph.D. dissertation, Dep. Elec. Eng. and Comput. Sci., Univ. California, Davis, Sept. 1982.
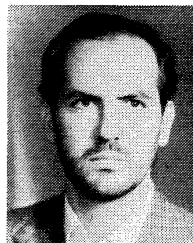
Mehrdad Hajivandi was born in Shiraz, Iran, on January 21, 1952. He received the B.S. degree from Shiraz University, Iran, in 1976, the M.S. degree from the University of California, Berkeley, in 1978, and the Ph.D. degree from the University of California, Davis, in 1982, all in electrical engineering.

He has been a faculty member at the Sharif University of Technology, Tehran, Iran, since 1982, where he is an Assistant Professor of Electrical Engineering. His research interests are in the general area of signal processing, with primary emphasis on communication and control theory, stochastic systems theory, and estimation and detection.
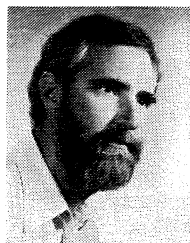
William A. Gardner (S'64–M'67–SM'84) was born in Palo Alto, CA, on November 4, 1942. He received the M.S. degree from Stanford University in 1967, and the Ph.D. degree from the University of Massachusetts, Amherst, in 1972, both in electrical engineering.

He was a Member of the Technical Staff at Bell Laboratories in Massachusetts from 1967 to 1969. He has been a faculty member at the University of California, Davis, since 1972, where he is Professor of Electrical Engineering and Computer Science. Since 1982, he has been President of the engineering consulting firm Statistical Signal Processing, Inc., Yountville, CA. His research interests are in the general area of statistical signal processing, with primary emphasis on the theories of time-series analysis, stochastic processes, and signal detection and estimation. He is the author of *Introduction to Random Processes with Applications to Signals and Systems* (Macmillan, 1985, second edition, McGraw-Hill, 1990), *The Random Processes Tutor: A Comprehensive Solutions Manual for Independent Study* (McGraw-Hill, 1990), and *Statistical Spectral Analysis: A Nonprobabilistic Theory* (Prentice-Hall, 1987). He holds several patents and is the author of over 50 research journal papers.

Dr. Gardner received the Best Paper of the Year Award from the European Association for Signal Processing in 1986, the 1987 Distinguished Engineering Alumnus Award from the University of Massachusetts, and the Stephen O. Rice Prize Paper Award in the Field of Communication Theory from the IEEE Communications Society in 1988. He is a member of the American Mathematical Society, the Mathematical Association of America, the American Association for the Advancement of Science, and the European Association for Signal Processing.