Tracking Analysis of the LMF and LMMN Adaptive Algorithms*

Nabil R. Yousef and Ali H. Sayed

Adaptive and Nonlinear Systems Laboratory
Electrical Engineering Department
University of California
Los Angeles, CA 90095

Abstract

Although the least mean fourth (LMF) and the least mean mixed norm (LMMN) adaptive algorithms are recommended for highly nonstationary environments, their tracking capabilities are not yet fully understood. This is mainly due to the fact that both algorithms involve nonlinear update equations for the weight error vector. In this paper we present a new approach to the tracking analysis of the LMF and LMMN algorithms, which bypasses the need for working directly with the weight error vector, and is based on a fundamental energy-preserving relation. By studying the energy flow through the system in steady-state, we derive expressions for the steady-state excess mean square error (EMSE) for both algorithms. We also derive optimal parameter values that minimize the EMSE in each case, and support our conclusions by simulations.

1 INTRODUCTION

Consider measurements $\{d(i)\}$ that arise from the linear model

$$d(i) = \mathbf{u}_i \mathbf{w}_i^o + v(i) \ , \tag{1}$$

where $v(i)$ accounts for measurement noise and modeling errors, and $\mathbf{u}_i$ denotes non zero *row* input (or regressor) vectors. Recursive estimates for the unknown weight vector $\mathbf{w}_i^o$ can be obtained adaptively via

$$\mathbf{w}_{i+1} = \mathbf{w}_i + \mu \, \mathbf{u}_i^* f_e(i) \ , \tag{2}$$

where $\mathbf{w}_i$ is an estimate for $\mathbf{w}_i^o$ at time i, μ is the algorithm step size, and $f_e(i)$ is a scalar function of the estimation error $e(i) = d(i) - \mathbf{u}_i \mathbf{w}_i$. The most popular variant of (2) is the least mean squares (LMS) algorithm [1], which corresponds to the linear error function $f_e^{LMS}(i) = e(i)$. Among other variants are the least mean fourth (LMF) [2] and the least mean mixed norm (LMMN) [3, 4] algorithms, which correspond to the error functions $f_e^{LMF}(i) = e^3(i)$ and $f_e^{LMMN}(i) = \delta \, e(i) + (1-\delta) \, e^3(i)$, respectively. The parameter δ is called the norm mixing parameter.

It is known that the LMF algorithm has better steady state performance than that of the LMS algorithm for applications in which the plant noise $v(i)$ has a probability density function with short tail [2]. However, its stability properties are worse than those of the LMS algorithm [3]. On the other hand, the LMMN algorithm has better steady state performance than the LMS algorithm and better stability properties than the LMF algorithm [4, 5].

Now, unlike the LMS algorithm, no tracking analysis is available in the literature for the LMF or LMMN algorithms. This is mainly due to the nonlinear nature of their update equations, which makes their tracking analysis using conventional approaches rather difficult. In this paper we propose a new approach to the tracking analysis of both algorithms. The approach bypasses the need for working directly with the weight error vector $\tilde{\mathbf{w}}_i = \mathbf{w}_i^o - \mathbf{w}_i$, and is based on a fundamental energy-preserving relation.

In a nonstationary environment, $\mathbf{w}_i^o$ is often assumed to vary randomly with time according to (see, e.g., [8]):

$$\mathbf{w}_{i+1}^o = \mathbf{w}_i^o + \mathbf{q}_i \ , \tag{3}$$

where $\mathbf{q}_i$ is a random vector. An important performance measure, which characterizes how well an adaptive algorithm tracks such variations in the environment, is the steady-state mean-square-error (MSE), defined by

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

Under the often realistic assumption that (see, *e.g.*, [1, 6, 7]):

*This material was based on work supported in part by the National Science Foundation under Award CCR-9732376 and ECS-9820765.

0-7803-5700-0/99/$10.00©1999 IEEE

A.1 The noise sequence $\{v(i)\}$ is identically individually distributed (iid) and statistically independent of the regressor sequence $\{\mathbf{u}_i\}$,

we find that the MSE is equivalently given by

$$\mathsf{MSE} = \sigma_v^2 + \lim_{i\to\infty} \mathrm{E}\,|\mathbf{u}_i\tilde{\mathbf{w}}_i|^2 \quad ,$$

which is dependent on $\tilde{\mathbf{w}}_i$. In the sequel we shall derive expressions for the steady-state excess mean square error (EMSE),

$$\zeta \triangleq \mathsf{MSE} - \sigma_v^2 = \lim_{i\to\infty} \mathrm{E}\,|\mathbf{u}_i\tilde{\mathbf{w}}_i|^2 \quad ,$$

as well as expressions for the optimum values of the step size and the norm mixing parameter that minimize the steady-state EMSE, ζ.

2 Fundamental Energy Relation

We start by noting that with any adaptive scheme we can associate the following so-called a-priori and a-posteriori estimation errors, $e_a(i) = \mathbf{u}_i\tilde{\mathbf{w}}_i$, $e_p(i) = \mathbf{u}_i(\tilde{\mathbf{w}}_{i+1} - \mathbf{q}_i)$. Using the data model (1), it is easy to see that the errors $\{e(i), e_a(i)\}$ are related via $e(i) = e_a(i) + v(i)$. If we further subtract $\mathbf{w}_i^o$ from both sides of (2) and multiply by $\mathbf{u}_i$ from the left, we also find that the errors $\{e_p(i), e_a(i), e(i)\}$ are related via:

$$e_p(i) = e_a(i) - \frac{\mu}{\bar{\mu}(i)} f_e(i) \,, \tag{4}$$

where we defined, for compactness of notation, $\bar{\mu}(i) = 1/\|\mathbf{u}_i\|^2$. Substituting (3) and (4) into (2), we obtain the update relation

$$\tilde{\mathbf{w}}_{i+1} = \tilde{\mathbf{w}}_i - \bar{\mu}(i)\mathbf{u}_i^*[e_a(i) - e_p(i)] + \mathbf{q}_i \ .$$

By evaluating the energies of both sides of this equation we obtain

$$\|\tilde{\mathbf{w}}_{i+1} - \mathbf{q}_i\|^2 + \bar{\mu}(i)|e_a(i)|^2 = \|\tilde{\mathbf{w}}_i\|^2 + \bar{\mu}(i)|e_p(i)|^2 \tag{5}$$

This energy conservation relation, first noted in [9]–[11], holds for all adaptive algorithms whose recursions are of the form given by (2); 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. It also establishes that the mapping from $\left\{\tilde{\mathbf{w}}_i, \sqrt{\bar{\mu}(i)}e_p(i)\right\}$ to $\left\{\tilde{\mathbf{w}}_{i+1} - \mathbf{q}_i, \sqrt{\bar{\mu}(i)}e_a(i)\right\}$ is energy preserving (or lossless). Furthermore, by combining (5) with (4), we see that both relations establish the existence of the feedback configuration shown in Figure 1, where $\mathcal{T}$ denotes a lossless map and q^{-1} denotes the unit delay operator. It could be seen from the figure that the system nonstationarity vector $\mathbf{q}_i$ acts as a disturbance input to the system.

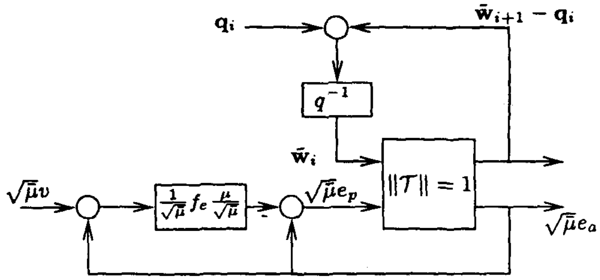

Figure 1. *Lossless mapping and a feedback loop.*

We now use the energy relation (5) to evaluate the MSE of an adaptive filter once it reaches steady-state. To do so, we impose the following assumptions, which are typical in the context of the tracking analysis of adaptive filters (see, *e.g.*, [8]).

A.2 The sequences $\{\mathbf{u}_i\}$ and $\{v(i)\}$ are mutually statistically independent of $\{\mathbf{q}_i\}$.

A.3 The sequence $\{\mathbf{q}_i\}$ is a stationary sequence of independent zero-mean vectors whose autocorrelation matrix $\mathbf{Q} = \mathrm{E}\,\mathbf{q}_i\mathbf{q}_i^*$ is positive definite.

Using $\mathrm{E}\,\|\tilde{\mathbf{w}}_{i+1}\|^2 = \mathrm{E}\,\|\tilde{\mathbf{w}}_i\|^2$ in steady-state, (4), A.2 and A.3, it is straightforward to verify that the energy relation (5) becomes

$$\mathrm{E}\,\bar{\mu}(i)|e_a(i)|^2 = \mathrm{Tr}(\mathbf{Q}) + \mathrm{E}\,\bar{\mu}(i)\left|e_a(i) - \frac{\mu}{\bar{\mu}(i)} f_e(i)\right|^2 \ . \tag{6}$$

We now show how to use this relation in evaluating the tracking performance of the LMF and LMMN algorithms.

3 Tracking Analysis

Throughout the analysis, we will only study the steady state performance of the LMMN algorithm and then treat the LMF algorithm as a special case (that corresponds to the choice $\delta = 0$). Now in steady state, and when μ is small enough, it is reasonable to assume that $|e_a(i)|^2 << |v(i)|^2$ (see [5]). Using $e(i) = e_a(i) + v(i)$, we can then write the error function of the LMMN algorithm as

$$f_e(i) \approx \delta[e_a(i) + v(i)] + (1-\delta)[3v^2(i)e_a(i) + v^3(i)] \ .$$

Introduce, for compactness of notation,

$$\bar{\delta} = 1 - \delta, \ \ \mathrm{E}\,|v(i)|^4 = \xi_v^4 \,, \ \ \mathrm{E}\,|v(i)|^6 = \xi_v^6 \ .$$

Using A.1 and the above approximation for $f_e(i)$, it is straightforward to show that equation (6) becomes

$$2\mu(\delta + 3\bar{\delta}\sigma_v^2)\zeta = \mathrm{Tr}(\mathbf{Q}) + \mu^2\,\mathrm{Tr}(\mathbf{R})\left(\delta^2\sigma_v^2 + 2\delta\bar{\delta}\xi_v^4 + \bar{\delta}^2\xi_v^6\right)$$
$$+\mu^2\,\mathrm{E}\,\|\mathbf{u}_i\|^2|e_a(i)|^2[\delta^2 + 6\delta\bar{\delta}\sigma_v^2 + 9\bar{\delta}^2\xi_v^4] \qquad (7)$$

To solve for ζ^{LMMN} we consider two cases:

1. For sufficiently small μ, we can assume that the third term on the RHS of (7) is negligible with respect to the second term, so that

$$\zeta^{\mathsf{LMMN}} = \frac{\mathrm{Tr}(\mathbf{Q})/\mu + \mu\,\mathrm{Tr}(\mathbf{R})\left(\delta^2\sigma_v^2 + 2\delta\bar{\delta}\xi_v^4 + \bar{\delta}^2\xi_v^6\right)}{2\left(\delta + 3\bar{\delta}\sigma_v^2\right)} \qquad (8)$$

At $\delta = 0$, equation (8) reduces to the EMSE of the LMF algorithm, which is given by

$$\zeta^{\mathsf{LMF}} = \frac{\mathrm{Tr}(\mathbf{Q})/\mu + \mu\,\mathrm{Tr}(\mathbf{R})\xi_v^6}{6\sigma_v^2} \ . \qquad (9)$$

2. For larger values of μ, equation (7) can be solved by imposing the following (often studied) assumption:

A.4 At steady state, $\mu^2\|\mathbf{u}_i\|^2$ is statistically independent of $|e_a(i)|^2$.

This assumption in fact becomes realistic for long filter lengths. Furthermore, it becomes *exact* for the case of constant modulus data that arises in some adaptive filtering applications [12, 13]. Using A.4, and (7), we then obtain

$$\zeta^{\mathsf{LMMN}} = \frac{\mathrm{Tr}(\mathbf{Q})/\mu + \mu\,\mathrm{Tr}(\mathbf{R})\left(\delta^2\sigma_v^2 + 2\delta\bar{\delta}\xi_v^4 + \bar{\delta}^2\xi_v^6\right)}{2\left(\delta + 3\bar{\delta}\sigma_v^2\right) - \mu\,\mathrm{Tr}(\mathbf{R})\left(\delta^2 + 6\delta\bar{\delta}\sigma_v^2 + 9\bar{\delta}\xi_v^4\right)} \qquad (10)$$

and

$$\zeta^{\mathsf{LMF}} = \frac{\mathrm{Tr}(\mathbf{Q})/\mu + \mu\,\mathrm{Tr}(\mathbf{R})\xi_v^6}{6\sigma_v^2 - 9\mu\,\mathrm{Tr}(\mathbf{R})\xi_v^4} \ . \qquad (11)$$

4 Parameter Optimization

We now investigate the existence of optimum design parameters $\{\delta_o, \mu_o\}$ that minimize the steady state EMSE for the LMMN and LMF algorithms, as given by (8) and (9). This is done for the following two cases:

A. Fixed δ and optimal μ

If the norm mixing parameter δ is a priori chosen to fullfill some convergence properties, then there will always exist an optimum value of μ that minimizes ζ^{LMMN}, which is directly given from (8) by

$$\mu_o^{\mathsf{LMMN}} = \sqrt{\mathrm{Tr}(\mathbf{Q})/\left(\mathrm{Tr}(\mathbf{R})\left(\delta^2\sigma_v^2 + 2\delta\bar{\delta}\xi_v^4 + \bar{\delta}^2\xi_v^6\right)\right)} \ . \qquad (12)$$

The corresponding minimum value of the steady state EMSE is given by

$$\zeta_{min}^{\mathsf{LMMN}} = \frac{\sqrt{\mathrm{Tr}(\mathbf{Q})\,\mathrm{Tr}(\mathbf{R})\left(\delta^2\sigma_v^2 + 2\delta\bar{\delta}\xi_v^4 + \bar{\delta}^2\xi_v^6\right)}}{\left(\delta + 3\bar{\delta}\sigma_v^2\right)} \ . \qquad (13)$$

The LMF algorithm always has a constrained δ that is equal to zero. Therefore, the optimum step size that minimizes its steady state EMSE, given in (9) and the corresponding minimum steady state EMSE, are respectively given by

$$\mu_o^{\mathsf{LMF}} = \sqrt{\mathrm{Tr}(\mathbf{Q})/\left(\mathrm{Tr}(\mathbf{R})\xi_v^6\right)} \ , \qquad (14)$$

and

$$\zeta_{min}^{\mathsf{LMF}} = \frac{\sqrt{\mathrm{Tr}(\mathbf{Q})\,\mathrm{Tr}(\mathbf{R})\xi_v^6}}{3\sigma_v^2} \ . \qquad (15)$$

B. Optimal δ and μ

Differentiating (8) separately with respect to μ and δ, we obtain

$$\frac{\partial\zeta}{\partial\mu} = \frac{-\mu^{-2}\,\mathrm{Tr}(\mathbf{Q}) + \mathrm{Tr}(\mathbf{R})(A\delta^2 + B\delta + C)}{2(D\delta + E)} \ , \qquad (16)$$

and

$$\frac{\partial\zeta}{\partial\delta} = \frac{1}{2(D\delta + E)^2}\left[\mu\,\mathrm{Tr}(\mathbf{R})(D\delta + E)(2A\delta + B)\right.$$
$$\left.-D\left(\mathrm{Tr}(\mathbf{Q})/\mu + \mu\,\mathrm{Tr}(\mathbf{R})(A\delta^2 + B\delta + C)\right)\right] \ , \qquad (17)$$

where $A = \left(\sigma_v^2 - 2\xi_v^4 + \xi_v^6\right)$, $B = 2\left(\xi_v^4 - \xi_v^6\right)$, $C = \xi_v^6$, $D = (1 - 3\sigma_v^2)$ and $E = 3\sigma_v^2$. Equating $\frac{\partial\zeta}{\partial\mu}$ and $\frac{\partial\zeta}{\partial\delta}$ to zero, it is straightforward to show that ζ has one local minimum or maximum at

$$\delta_o = \frac{2DC - BE}{2AE - BD} \ , \qquad (18)$$

and

$$\mu_o = \sqrt{\mathrm{Tr}(\mathbf{Q})/\left(\mathrm{Tr}(\mathbf{R})(A\delta_o^2 + B\delta_o + C)\right)} \ . \qquad (19)$$

The pair $\{\delta_o, \mu_o\}$ corresponds to a minimum iff ζ possesses a positive definite Hessian matrix $\mathbf{H}(\delta, \mu)$, which is defined by

$$\mathbf{H}(\delta,\mu) = \begin{bmatrix} \frac{\partial^2\zeta}{\partial\mu^2} & \frac{\partial^2\zeta}{\partial\delta\partial\mu} \\ \frac{\partial^2\zeta}{\partial\mu\partial\delta} & \frac{\partial^2\zeta}{\partial\delta^2} \end{bmatrix} \ . \qquad (20)$$

Differentiating (16) and (17) to obtain the second partial derivatives of ζ and substituting into (20), it can be shown that the Hessian matrix at $\{\delta_o, \mu_o\}$ is given by

$$\mathbf{H}(\delta_o,\mu_o) = \begin{bmatrix} \frac{\mathrm{Tr}(\mathbf{Q})}{\mu_o^3(D\delta_o+E)} & \frac{D\,\mathrm{Tr}(\mathbf{Q})}{\mu_o(D\delta_o+E)} \\ \frac{D\,\mathrm{Tr}(\mathbf{Q})}{\mu_o(D\delta_o+E)} & \frac{\mu_o A\,\mathrm{Tr}(\mathbf{R})}{(D\delta_o+E)} \end{bmatrix} \ .$$

The eigenvalues of $\mathbf{H}(\delta_o, \mu_o)$ are the solutions of the characteristic equation

$$\begin{vmatrix} \lambda - \frac{\mathrm{Tr}(\mathbf{Q})}{\mu_o^3(D\delta_o+E)} & \frac{D\,\mathrm{Tr}(\mathbf{Q})}{\mu_o(D\delta_o+E)} \\ \frac{D\,\mathrm{Tr}(\mathbf{Q})}{\mu_o(D\delta_o+E)} & \lambda - \frac{\mu_o A\,\mathrm{Tr}(\mathbf{R})}{(D\delta_o+E)} \end{vmatrix} = 0 \ ,$$

that is

$$\lambda_{1,2} = \frac{\gamma}{2} \pm \frac{1}{2}\sqrt{\gamma^2 + 4\left(\frac{D^2\,\mathrm{Tr}^2(\mathbf{Q}) - A\,\mathrm{Tr}(\mathbf{Q})\,\mathrm{Tr}(\mathbf{R})}{\mu_o^2(D\delta_o + E)^2}\right)} ,$$

where

$$\gamma = \frac{\mu_o^{-3}\,\mathrm{Tr}(\mathbf{Q}) + A\mu_o\,\mathrm{Tr}(\mathbf{R})}{D\delta_o + E} .$$

For the two eigenvalues to be positive, the system parameters should satisfy $A > 0$ and $D^2\,\mathrm{Tr}(\mathbf{Q}) < A \cdot \mathrm{Tr}(R)$. Furthermore, since $\delta_o \in [0, 1]$, the term $(D\delta_o + E)$ is strictly positive. Using (8), we deduce that if the system noise moments and the degree of nonstationarity satisfy

$$\sigma_v^2 + \xi_v^6 > 2\xi_v^4 , \tag{21}$$

and

$$\mathrm{Tr}(\mathbf{Q}) < \frac{\mathrm{Tr}(\mathbf{R})(\sigma_v^2 - 2\xi_v^4 + \xi_v^6)}{(1 - 3\sigma_v^2)^2} , \tag{22}$$

then the steady state EMSE of the LMMN algorithm will have a well defined minimum at $\{\delta_o, \mu_o\}$ given respectively by

$$\delta_o = \frac{(1 - 3\sigma_v^2)^2\xi_v^6 - 3\sigma_v^2\left(\xi_v^4 - \xi_v^6\right)}{3\sigma_v^2\left(\sigma_v^2 - 2\xi_v^4 + \xi_v^6\right) - (1 - 3\sigma_v^2)^2\left(\xi_v^4 - \xi_v^6\right)} , \tag{23}$$

and

$$\mu_o = \sqrt{\mathrm{Tr}(\mathbf{Q}) / \left(\mathrm{Tr}(\mathbf{R})\left(\delta_o^2\sigma_v^2 + 2\delta_o(\bar{\delta}_o)\xi_v^4 + (\bar{\delta}_o)^2\xi_v^6\right)\right)} \tag{24}$$

Moreover, the corresponding minimum value of the EMSE will be

$$\zeta_{min}^{\mathsf{LMMN}} = \frac{\sqrt{\mathrm{Tr}(\mathbf{Q})\,\mathrm{Tr}(\mathbf{R})[\delta_o^2\sigma_v^2 + 2\delta_o(\bar{\delta}_o)\xi_v^4 + (\bar{\delta}_o)^2\xi_v^6]}}{\delta_o + 3(\bar{\delta}_o)\sigma_v^2} . \tag{25}$$

Gaussian Plant Noise

For Gaussian system noise, $\xi_v^4 = 3\sigma_v^4$ and $\xi_v^6 = 15\sigma_v^6$. Then

$$\begin{aligned} \sigma_v^2 + \xi_v^6 - 2\xi_v^4 &= \sigma_v^2(1 - 6\sigma_v^2 + 15\sigma_v^4) \\ &= 15\sigma_v^2\left(\left(\sigma_v^2 - \frac{1}{5}\right)^2 + \frac{2}{75}\right) > 0 , \end{aligned}$$

which implies that (21) is always true for the Gaussian noise case. Then, if the system degree of nonstationarity satisfies

$$\mathrm{Tr}(\mathbf{Q}) < \frac{\sigma_v^2(1 - 6\sigma_v^2 + 15\sigma_v^4)\,\mathrm{Tr}(\mathbf{R})}{(1 - 3\sigma_v^2)^2} , \tag{26}$$

the optimum value of δ is given from (23) by

$$\delta_o = \frac{15\sigma_v^6(1 - 3\sigma_v^2)^2 - 3\sigma_v^2\left(3\sigma_v^4 - 15\sigma_v^6\right)}{3\sigma_v^2\left(\sigma_v^2 - 6\sigma_v^4 + 15\sigma_v^6\right) - (1 - 3\sigma_v^2)^2\left(3\sigma_v^4 - 15\sigma_v^6\right)} = 1,$$

which corresponds to the LMS algorithm with an optimal step size given by

$$\mu_o = \sqrt{\frac{\mathrm{Tr}(\mathbf{Q})}{\sigma_v^2\,\mathrm{Tr}(\mathbf{R})}} ,$$

and a corresponding minimum EMSE of

$$\zeta_{min} = \sigma_v\sqrt{\mathrm{Tr}(\mathbf{Q})\,\mathrm{Tr}(\mathbf{R})} .$$

That is, for Gaussian system noise, if (26) holds, the LMS algorithm outperforms the LMF and LMMN algorithms; an interesting conclusion.

5 Simulation Results

Several simulations are carried out to validate our theoretical results. In the simulations, the unknown system weight vector $\mathbf{w}_i^o$ is of length 10 and the elements of the system nonstationarity and input vectors, $\mathbf{q}_i$ and $\mathbf{u}_i$, are white Gaussian of variances σ_q^2 and unity, respectively. The plant noise is chosen to be a linear combination of normally and uniformly distributed independent random variables of variances σ_n^2 and σ_u^2, respectively. Each simulation result is the steady state statistical average of 100 runs, with 10^5 iterations in each run.

Figure 2 compares the simulation and theoretical results of the steady state MSE of the LMMN algorithm for $\sigma_q = 5 \times 10^{-4}$, $\delta = 0.8$, $\sigma_n^2 = 10^{-2}$, and $\sigma_u^2 = 10^{-2}/12$. It is seen in the figure that the theoretical and experimental MSE are in good match. The figure also shows that the steady state MSE possesses a minimum value of 0.0113 at $\mu = 0.006$, which are in good agreement with the corresponding theoretical values obtained from expressions (13) and (12) as 0.01136 and $\mu_o = 0.0061$, respectively.

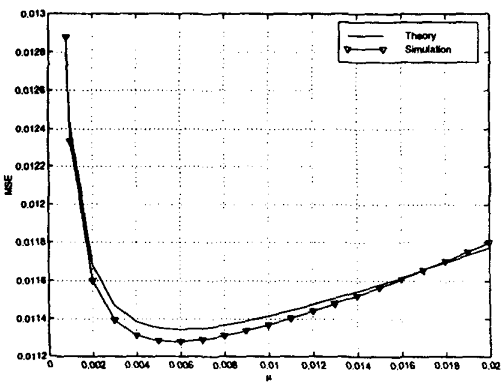

Figure 2. *Theory and simulation* MSE *vs.* μ.

Figure 3 compares the theoretical MSE obtained from expressions (8) and (10) with the experimental MSE for $\sigma_q = 10^{-5}$, $\delta = 0.5$, $\sigma_n^2 = 10^{-6}$, and $\sigma_u^2 = 10^{-4}/12$. The figure shows that both expressions are in good match with simulation results at small values of μ. However, expression (10) is in a much better match with the simulation results for relatively larger values of μ, which validates the use of assumption A.4.

Figure 4 shows the experimental MSE and the theoretical MSE obtained from expression (8) versus the norm mixing parameter δ for Gaussian plant noise of variance $\sigma_n^2 = 10^{-2}$, $\sigma_q = 10^{-3}$, and $\mu = 0.001$. It is clear that the minimum value of the MSE occurs at $\delta = 1$ for Gaussian noise.

Figure 5 shows the theoretical and simulated EMSE versus μ, for the optimal value of δ calculated from ex-

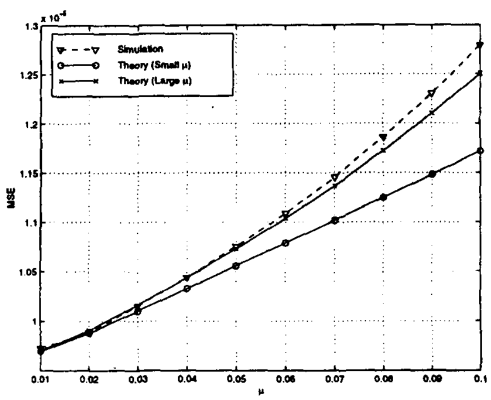

Figure 3. MSE *vs.* μ *for* $\delta = 0.5$. *Small and large* μ *theoretical* MSE *correspond to (8) and (10), respectively.*

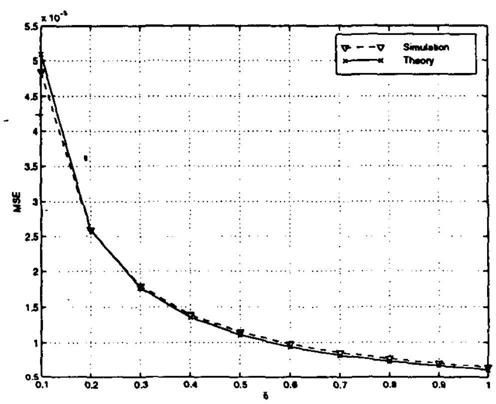

Figure 4. *Theory and simulation* MSE *vs.* δ *for Gaussian plant noise.*

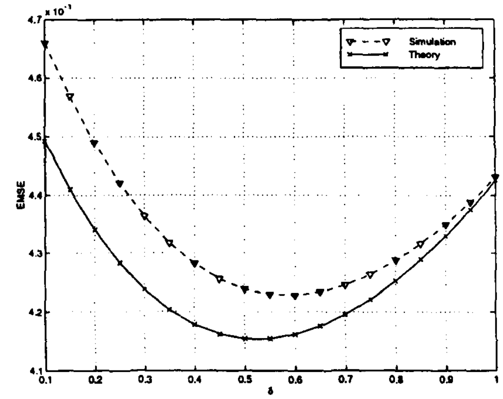

Figure 5. *Theory and simulation* EMSE *vs.* δ *at* $\mu = \mu_o$.

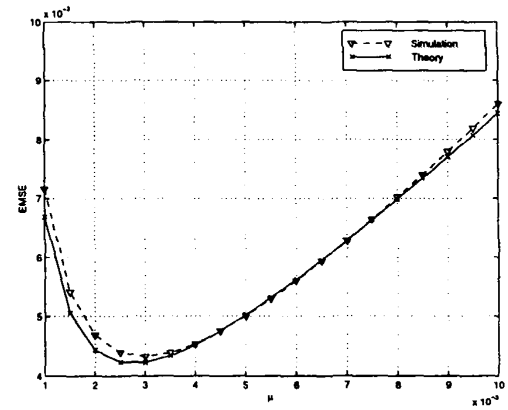

Figure 6. *Theory and simulation* EMSE *vs.* μ *at* $\delta = \delta_o$.

pression (23) to be $\delta_o = 0.5432$, $\sigma_n^2 = \sigma_u^2 = 0.1$, and $\sigma_q = 10^{-3}$. Figure 6 shows theoretical and simulated results versus δ, for the optimal value of μ calculated from expression (24) to be $\mu_o = 0.0029$. Both simulations show that optimal parameter values obtained from simulations, $\{\delta_o, \mu_o\} = \{0.59, 0.003\}$, are in good match with the values, $\{0.5432, 0.0029\}$, given by (23) and (24), respectively.

References

[1] B. Widrow and S. D. Stearns. *Adaptive Signal Processing*, Prentice Hall, NJ, 1985.

[2] E. Walach and B. Widrow. The least mean fourth (LMF) adaptive algorithm and its family. *IEEE Transactions on Information Theory*, vol. 30, no. 2, March 1994.

[3] J.A. Chambers, O. Tanrikulu and A.G. Constantindes. Least mean mixed-norm adaptive filtering. *Electronics Letters*, vol. 30, no. 19, September 1994.

[4] D.I. Pazaitis and A.G. Constantindes. LMS+F algorithm. *Electronics Letters*, vol. 31, no. 17, August 1995.

[5] O. Tanrikulu and J.A. Chambers. Convergence and steady-state properties of the least-mean mixed-norm (LMMN) adaptive algorithm. *IEE Proceedings-Vision, Image and Signal Processing*, vol. 143, no. 3, June 1996.

[6] S. Haykin. *Adaptive Filter Theory*. Prentice Hall, 3rd edition, NJ, 1996.

[7] O. Macchi. *Adaptive Processing: The LMS Approach with Applications in Transmission*. Wiley, NY, 1995.

[8] E. Eweda. Comparison of RLS, LMS, and sign algorithms for tracking randomly time-varying channels. *IEEE Transactions on Signal Processing*, 42(11), November 1994.

[9] A. H. Sayed and M. Rupp. A time-domain feedback analysis of adaptive algorithms via the small gain theorem. *Proc. SPIE*, vol. 2563, pp. 458–69, San Diego, CA, Jul. 1995.

[10] M. Rupp and A. H. Sayed. A time-domain feedback analysis of filtered-error adaptive gradient algorithms. *IEEE Transactions on Signal Processing*, vol. 44, no. 6, pp. 1428–1439, Jun. 1996.

[11] A. H. Sayed and M. Rupp. Robustness issues in adaptive filtering. *DSP Handbook*, Chapter 20, CRC Press, 1998.

[12] N. R. Yousef and A. H. Sayed. A unified approach to the steady-state and tracking analyses of adaptive filtering algorithms. *Proc. of IEEE-EURASIP Workshop on Nonlinear Signal and Image Processing*, vol. 2, pp. 699-703, Antalya, Turkey, June 1999.

[13] N. R. Yousef and A. H. Sayed. A feedback analysis of the tracking performance of blind adaptive equalization algorithms. *Proc. of 1999 IEEE Conference on Decision and Control*, Phoenix, AZ, Dec. 1999.