A GENERALIZED TRACKING ANALYSIS OF ADAPTIVE FILTERING ALGORITHMS IN CYCLICLY AND RANDOMLY VARYING ENVIRONMENTS

NABIL R. YOUSEF AND ALI H. SAYED

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

ABSTRACT

This paper derives new tracking results for adaptive filtering algorithms operating in the presence of two sources of nonstationarities: a carrier frequency offset and random variations. Both impairments are common in digital communications due to variations in channel characteristics and to mismatches between transmitter and receiver carrier generators.

1. INTRODUCTION

Cyclic system nonstationarities arise in communication systems due to mismatches between the transmitter and receiver carrier generators. The ability of adaptive filtering algorithms to track such system variations are not fully understood. A recent contribution in this regard is the work [1], which performed a first-order analysis of the performance of the LMS algorithm in the presence of a carrier frequency offset.

In this paper, we develop a general framework for the tracking analysis of adaptive algorithms that can handle both cyclic as well as random system nonstationarities simultaneously. The framework is based on a fundamental variance (conservation) relation and it allows us to derive several new tracking results, as well as optimum design parameters, for several adaptive filtering algorithms (e.g., LMS, NLMS, LMF, LMMN, and Sign algorithms). In so doing, we also obtain expressions for the excess mean-square error in steady-state for all these algorithms.

1.1. The Model

Consider noisy measurements $\{d(i)\}$ that arise from a model of the form

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

where $v(i)$ accounts for measurement noise and modeling errors, $\mathbf{u}_i$ denotes a nonzero *row* input (regressor) vector, and $\mathbf{w}_i^o$ is an unknown *column* vector that we wish to track. The multiplicative term $e^{j\Omega i}$ accounts for a possible frequency offset between the transmitter and receiver carriers in a digital communications scenario. Furthermore, we assume that the unknown system vector $\mathbf{w}_i^o$ is randomly changing according to

$$\mathbf{w}_i^o = \mathbf{w}^o + \bar{\mathbf{q}}_i, \tag{2}$$

where $\mathbf{w}^o$ is a fixed vector, and where $\bar{\mathbf{q}}_i$ is assumed to be a zero-mean stationary random vector process with a positive-definite covariance matrix $\bar{\mathbf{Q}}$. It is also statistically independent of the sequences $\{v(i)\}$ and $\{\mathbf{u}_i\}$.

We thus see that the generalized system model, given by (1) and (2), includes the effects of both cyclic and random system nonstationarities; both of which are common impairments in communications systems and especially in applications that involve channel estimation, channel equalization, and inter-symbol-interference cancellation.

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

Table 1: *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	$\text{sign}[e(i)]$

1.2. Algorithms

The main purpose of this paper is to study the ability of adaptive filtering algorithms to estimate and track such cyclic and random variations in $\mathbf{w}_i^o$. We consider general adaptive schemes of the form

$$\mathbf{w}_{i+1} \quad = \quad \mathbf{w}_i + \mu \, \mathbf{u}_i^* \, f_e(i) \ , \tag{3}$$

where $\mathbf{w}_i$ is an estimate for $\mathbf{w}_i^o$ at iteration i, μ is the step-size, and $f_e(i)$ denotes a generic scalar function of the so-called output estimation error, defined by

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

Different choices for $f_e(i)$ result in different adaptive algorithms. Table 1 defines $f_e(i)$ for many famous special cases of (3) — see [2, 3].[1]

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

0-7803-6293-4/00/$10.00©2000 IEEE

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

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

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

$$\tilde{\mathbf{w}}_i = \mathbf{w}_i^o e^{j\Omega i} - \mathbf{w}_i \ . \tag{4}$$

Under the often realistic assumption that:

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

we find that the MSE is equivalently given by

$$\mathsf{MSE} = \sigma_v^2 + \lim_{i\to\infty} \mathrm{E}\,|\mathbf{u}_i\tilde{\mathbf{w}}_i|^2 \ . \tag{5}$$

We now proceed to derive expressions for the steady-state excess mean-square-error (EMSE), $\zeta = \lim_{i\to\infty} \mathrm{E}\,|\mathbf{u}_i\tilde{\mathbf{w}}_i|^2$, for various algorithms, along with values for the optimum algorithm parameters that minimize the EMSE.

2. FUNDAMENTAL ENERGY RELATION

Using (2) and (3), we obtain the following recursion for the weight-error vector

$$\tilde{\mathbf{w}}_{i+1} = \tilde{\mathbf{w}}_i - \mu\, \mathbf{u}_i^*\, f_e(i) + \mathbf{c}_i e^{j\Omega i} \ , \tag{6}$$

where $\mathbf{c}_i$ is defined by

$$\mathbf{c}_i \triangleq \mathbf{w}^o(e^{j\Omega} - 1) + \bar{\mathbf{q}}_{i+1}e^{j\Omega} - \bar{\mathbf{q}}_i. \tag{7}$$

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

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

Using the data model (1), it is then easy to see that $e(i) = e_a(i) + v(i)$. Moreover, if we further multiply (6) by $\mathbf{u}_i$ from the left, we also find that

$$e_p(i) = e_a(i) - \frac{\mu}{\bar{\mu}(i)} f_e(i) \ , \tag{8}$$

where we defined, for compactness, $\bar{\mu}(i) = 1/\|\mathbf{u}_i\|^2$. Substituting (8) into (6), we obtain the update relation

$$\tilde{\mathbf{w}}_{i+1} = \tilde{\mathbf{w}}_i - \bar{\mu}(i)\mathbf{u}_i^*[e_a(i) - e_p(i)] + \mathbf{c}_i e^{j\Omega i} \ . \tag{9}$$

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

$$\|\tilde{\mathbf{w}}_{i+1} - \mathbf{c}_i e^{j\Omega i}\|^2 + \bar{\mu}(i)|e_a(i)|^2 = \|\tilde{\mathbf{w}}_i\|^2 + \bar{\mu}(i)|e_p(i)|^2 \tag{10}$$

This energy conservation relation, first noted in [4, 5], holds for all adaptive algorithms whose recursions are of the form given by (3); 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.

2.1. Relevance to the Tracking Analysis

We now use the energy relation (10) to evaluate the EMSE of an adaptive filter once it reaches steady-state. To do so, we make the following reasonable assumption (see, e.g., [1]):

A.2 In steady-state, the weight-error vector $\tilde{\mathbf{w}}_i$ takes the generic form $\mathbf{z}_i e^{j\Omega i}$, with the stationary random process $\mathbf{z}_i$ independent of the offset frequency Ω. Let $\mathbf{z}$ denote $\mathrm{E}\,\mathbf{z}_i$.

Using (8), A.2, and $\mathrm{E}\,\|\tilde{\mathbf{w}}_{i+1}\|^2 = \mathrm{E}\,\|\tilde{\mathbf{w}}_i\|^2$ in steady-state, and taking expectations of both sides of (10), it can be verified that

$$\begin{aligned}\mathrm{E}\,\bar{\mu}(i)|e_a(i)|^2 &= 2\,\mathrm{Tr}(\bar{\mathbf{Q}}) + \|\mathbf{w}^o\|^2\,|1-e^{j\Omega}|^2 \\ &\quad -2\,\mathrm{Re}\,\mathrm{E}\left[\bar{\mathbf{q}}_i^*\left(\mathbf{z}_i - \mu\mathbf{u}_i^* f_e(i)e^{-j\Omega i}\right)\right] \\ &\quad -2\,\mathrm{Re}\left[(1-e^{j\Omega})^*\mathbf{w}^{o*}\,\mathrm{E}\left(\mathbf{z}_i - \mu\mathbf{u}_i^* f_e(i)e^{-j\Omega i}\right)\right] \\ &\quad +\mathrm{E}\,\bar{\mu}(i)\left|e_a(i) - \frac{\mu}{\bar{\mu}(i)} f_e(i)\right|^2 ,\end{aligned} \tag{11}$$

where $\mathrm{Tr}(\bar{\mathbf{Q}}) = \mathrm{E}\,\bar{\mathbf{q}}_i\bar{\mathbf{q}}_i^*$. This equation can now be solved for the steady-state excess mean-square-error (EMSE). Before proceeding with solving (11) for ζ, we can draw the following important conclusions:

1. The effect of the random system nonstationarity is represented by the first and third terms on the RHS of (11). The first term, $2\,\mathrm{Tr}(\bar{\mathbf{Q}})$, is independent of the frequency offset Ω and of the particular adaptive algorithm that is being employed. The third term is dependent on the algorithm and independent of Ω. Cyclic system nonstationarities contribute to the second and fourth terms on the RHS of (11). The second term, $\|\mathbf{w}^o\|^2|1-e^{j\Omega}|^2$, depends only on Ω and $\|\mathbf{w}^o\|^2$. However, the fourth term depends on the algorithm error function, $f_e(i)$.
2. The fourth term on the RHS of (11) could be evaluated in the following manner. First, we multiply by $e^{-j\Omega i}$ and apply the expectation operator to both sides of (6) to get

   $$(1-e^{j\Omega})\mathbf{z} = \mu\,\mathrm{E}\left(\mathbf{u}_i^* f_e(i)e^{-j\Omega i}\right) + \mathbf{w}^o(1-e^{j\Omega}). \tag{12}$$

   Second, we solve the above equation for $\mathbf{z}$ at steady-state. A similar procedure can be used to evaluate the third term on the RHS of (11). Several examples for using this procedure are given in the next section for various adaptive algorithms.
3. We may also mention that it can be verified that the LHS and the last term on the RHS of (11) are independent of Ω for all the algorithms listed in Table I. Thus, these terms can be set to their values in the stationary case (i.e., at Ω=0).

3. TRACKING ANALYSIS

We now apply the above general procedure to various adaptive algorithms from Table I. Due to space limitations, we omit some of the details and only highlight the main steps in the arguments. The reader will soon realize the convenience of working with (11) — see, e.g., [6, 7] for other steady-state and tracking results in the absence of cyclic nonstationarities and for random variations that are modeled by a first-order difference equation as opposed to (2).

3.1. The LMS Algorithm

For LMS we have $f_e(i) = e(i) = e_a(i) + v(i)$. First, to calculate $\mathrm{E}\,\mathbf{z}_i$ at steady-state, we impose the widely used independence assumption [3]:

A.3 At steady state, $\tilde{\mathbf{w}}_i$ is statistically independent of $\mathbf{u}_i$.

Substituting $f_e(i)$ into (12) and using A.1, A.2 and A.3, it follows immediately that (see [1])

$$\mathbf{z} = \mathrm{E}\,\mathbf{z}_i = \left[\mathbf{I} - \frac{\mu \mathbf{R}}{1 - e^{j\Omega}}\right]^{-1} \mathbf{w}^o , \tag{13}$$

where $\mathbf{R} = \mathrm{E}\,\mathbf{u}_i \mathbf{u}_i^*$. Substituting into (11) and using A.1, we obtain

$$2\mu\zeta^{\mathsf{LMS}} = 2\mu\,\mathrm{Tr}(\bar{\mathbf{Q}}\mathbf{R}) + \mu^2\sigma_v^2\,\mathrm{Tr}(\mathbf{R}) + \mu^2\,\mathrm{E}\,\|\mathbf{u}_i\|^2\,|e_a(i)|^2 + |1 - e^{j\Omega}|^2\,\mathrm{Re}\,\mathrm{Tr}\left[\mathbf{W}^o\left(\mathbf{I} - 2\mathbf{X}\right)\right] , \tag{14}$$

where $\mathbf{W}^o$ and $\mathbf{X}$ are defined by $\mathbf{W}^o = \mathbf{w}^o\mathbf{w}^{o*}$ and

$$\mathbf{X} = (\mathbf{I} - \mu\mathbf{R})\left(\mathbf{I} - \mu\mathbf{R} - e^{j\Omega}\mathbf{I}\right)^{-1} .$$

To solve for ζ^{LMS} we consider three cases:

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

$$\zeta^{\mathsf{LMS}} = \frac{\mu}{2}\sigma_v^2\,\mathrm{Tr}(\mathbf{R}) + \mathrm{Tr}(\bar{\mathbf{Q}}\mathbf{R}) + \frac{\mu^{-1}}{2}\beta, \tag{15}$$

where

$$\beta = |1 - e^{j\Omega}|^2\,\mathrm{Re}\,\mathrm{Tr}\left[\mathbf{W}^o\left(\mathbf{I} - 2\mathbf{X}\right)\right] . \tag{16}$$

2. For larger values of μ, equation (14) can be solved by imposing the following (often studied) assumption (which is realistic for long filter lengths):

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

Using A.2, and (14) we directly obtain

$$\zeta^{\mathsf{LMS}} = \frac{\mu\sigma_v^2\,\mathrm{Tr}(\mathbf{R}) + 2\,\mathrm{Tr}(\bar{\mathbf{Q}}\mathbf{R}) + \mu^{-1}\beta}{2 - \mu\,\mathrm{Tr}(\mathbf{R})} . \tag{17}$$

3. For Gaussian white-input signals ($\mathbf{R} = \sigma_u^2\mathbf{I}$), equation (14) can be more accurately solved by using A.3 to yield

$$\zeta^{\mathsf{LMS}} = \frac{\mu M\sigma_u^2\sigma_v^2 + 2\sigma_u^2\,\mathrm{Tr}(\bar{\mathbf{Q}}) + c}{2 - \mu(M + \lambda)\sigma_u^2} , \tag{18}$$

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,

$$c = \sigma_u^2\left(2 - \mu\sigma_u^2\right)\frac{|1 - e^{j\Omega}|^2}{|1 - \mu\sigma_u^2 - e^{j\Omega}|^2}\,\|\mathbf{w}^o\|^2 .$$

For small values of Ω and $\mu\sigma_u^2 \gg (1-\cos\Omega)$, which is usually valid in practical cases, this term can be approximated by

$$c \approx \frac{\Omega^2\left(2 - \mu\sigma_u^2\right)}{\mu^2\sigma_u^2}\,\|\mathbf{w}^o\|^2 . \tag{19}$$

From these results 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 three terms. The first term increases with μ, the noise variance σ_v^2, and σ_u^2. The second term is independent of μ and increases with the random nonstationarity term $\mathrm{Tr}(\bar{\mathbf{Q}})$. The third 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 in this case. This optimal value can be obtained by minimizing the EMSE in (18) over μ. A *rough* estimate for μ_o can be obtained, from (18) and (19), to be

$$\mu_o^{\mathsf{LMS}} \approx \left(\frac{4\,\Omega^2\,\|\mathbf{w}^o\|^2}{M\sigma_v^2\sigma_u^4}\right)^{1/3} . \tag{20}$$

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.

3.2. The LMF and LMMN Algorithms

For the case of the LMF and LMMN algorithms, we need only study the tracking performance of the LMMN algorithm and then obtain the LMF algorithm as a special case by setting $\delta = 0$. We assume the noise sequence is circular so that $\mathrm{E}\,v^2(i) = 0$. Introduce also, for compactness of notation,

$$\bar{\delta} = 1 - \delta,\ \mathrm{E}\,|v(i)|^4 = \xi_v^4\ ,\ \mathrm{E}\,|v(i)|^6 = \xi_v^6,\ \ \gamma = \delta + 2\bar{\delta}\sigma_v^2.$$

Now in steady-state, and when μ is small enough, it is reasonable to assume that $|e_a(i)|^2 \ll |v(i)|^2$. Using $e(i) = e_a(i) + v(i)$, we can then write the error function of the LMMN algorithm as

$$f_e \approx \delta[e_a + v] + \bar{\delta}[2e_a|v|^2 + v|v|^2 + v^2 e_a^*] .$$

Substituting $f_e(i)$ into (12) and using A.1, A.2, and A.4, yields

$$\mathbf{z} = \left[\mathbf{I} - \frac{\mu\gamma\mathbf{R}}{(1 - e^{j\Omega})}\right]^{-1} \mathbf{w}^o .$$

Substituting into (11) and using A.1, one obtains

$$2\mu\gamma\zeta = 2\mu\gamma\,\mathrm{Tr}(\bar{\mathbf{Q}}\mathbf{R}) + \mu^2 a\,\mathrm{Tr}(\mathbf{R}) + \mu^2 b\,\mathrm{E}\,\|\mathbf{u}_i\|^2|e_a(i)|^2 + |1 - e^{j\Omega}|^2\,\mathrm{Re}\,\mathrm{Tr}\left[\mathbf{W}^o\left(\mathbf{I} - 2\mathbf{X}\right)\right] , \tag{21}$$

where $a = \left(\delta^2\sigma_v^2 + 2\delta\bar{\delta}\xi_v^4 + \bar{\delta}^2\xi_v^6\right)$, $b = (\delta^2 + 8\delta\bar{\delta}\sigma_v^2 + 9\bar{\delta}^2\xi_v^4)$, and $\mathbf{X} = [\mathbf{I} - \mu\gamma\mathbf{R}]\left[\mathbf{I} - \mu\gamma\mathbf{R} - e^{j\Omega}\mathbf{I}\right]^{-1}$.

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

1. For sufficiently small μ, we can assume that the third term on the RHS of (21) is negligible, so that

$$\zeta^{\text{LMMN}} = \frac{\mu a \operatorname{Tr}(\mathbf{R}) + 2\gamma \operatorname{Tr}(\bar{\mathbf{Q}}\mathbf{R}) + \mu^{-1}\beta}{2\gamma}$$

where β is defined as in (16). At $\delta = 0$, the above equation reduces to the EMSE of the LMF algorithm, which is given by

$$\zeta^{\text{LMF}} = \frac{\mu \xi_v^6 \operatorname{Tr}(\mathbf{R}) + 4\sigma_v^2 \operatorname{Tr}(\bar{\mathbf{Q}}\mathbf{R}) + \mu^{-1}\beta}{4\sigma_v^2} ,$$

where $\mathbf{X}$, for the LMF algorithm, is given by

$$\mathbf{X} = \left(\mathbf{I} - 2\mu\sigma_v^2\mathbf{R}\right)\left(\mathbf{I} - 2\mu\sigma_v^2\mathbf{R} - e^{j\Omega}\mathbf{I}\right)^{-1} .$$

2. For larger values of μ, by imposing A.5, equation (21) leads to

$$\zeta^{\text{LMMN}} = \frac{\mu a \operatorname{Tr}(\mathbf{R}) + 2\gamma \operatorname{Tr}(\bar{\mathbf{Q}}\mathbf{R}) + \mu^{-1}\beta}{2\gamma - \mu b \operatorname{Tr}(\mathbf{R})}$$

3. For Gaussian white input signals with $\mathbf{R} = \sigma_u^2\mathbf{I}$, equation (21) can be solved by imposing A.4 to obtain

$$\zeta^{\text{LMMN}} = \frac{\mu a M \sigma_u^2 + 2\gamma\sigma_u^2 \operatorname{Tr}(\bar{\mathbf{Q}}) + c}{2\gamma - \mu b (M + \lambda)\sigma_u^2}$$

where c is given by

$$c = \gamma\sigma_u^2\left(2 - \mu\gamma\sigma_u^2\right)\frac{|1 - e^{j\Omega}|^2}{|1 - \mu\gamma\sigma_u^2 - e^{j\Omega}|^2}\,\|\mathbf{w}^o\|^2 ,$$

3.3. The Sign Algorithm

For the SA we have $f_e(i) = \text{sign}[e(i)] = \text{sign}[e_a(i) + v(i)]$. Using Price theorem [8], (11), (12), (16), A.1, A.2, and A.4, assuming Gaussian signals $\{\mathbf{u}_i, v(i), e(i)\}$, real-valued regressors $\{\mathbf{u}_i\}$, and using $|e_a(i)|^2 \ll \sigma_v^2$ in steady-state, it is straightforward to show that

$$\zeta^{\text{SA}} = \frac{\mu \operatorname{Tr}(\mathbf{R}) + 2\alpha \operatorname{Tr}(\bar{\mathbf{Q}}\mathbf{R}) + \mu^{-1}\beta}{2\alpha}$$

where $\alpha = \sqrt{\frac{2}{\pi\sigma_v^2}}$ and $\mathbf{X} = [\mathbf{I} - \mu\alpha\mathbf{R}]\left[\mathbf{I} - \mu\alpha\mathbf{R} - e^{j\Omega}\mathbf{I}\right]^{-1}$.

Finally, we may add that values for the optimum algorithm parameters could be evaluated by minimizing the EMSE expressions for the LMF, LMMN and sign algorithms using the same procedure used for the LMS algorithm in Section 3.1.

4. SIMULATION RESULTS

Figure 1 compares the theoretical and experimental MSE of the LMS algorithm for a wide range of the step-size μ and for three different values of the carrier offset Ω (0.0001, 0.0002, 0.0004). In the simulations, we used a white Gaussian input signal of unity variance, a 10 tap unknown system, $\sigma_v = 3 \times 10^{-2}$, $\operatorname{Tr}(\bar{Q}) = 10^{-7}$. Each simulation point is the average of 50 runs with 3000 iterations in each run.

It is clear from the figure that the theoretical results are in very good match with the simulation results. For

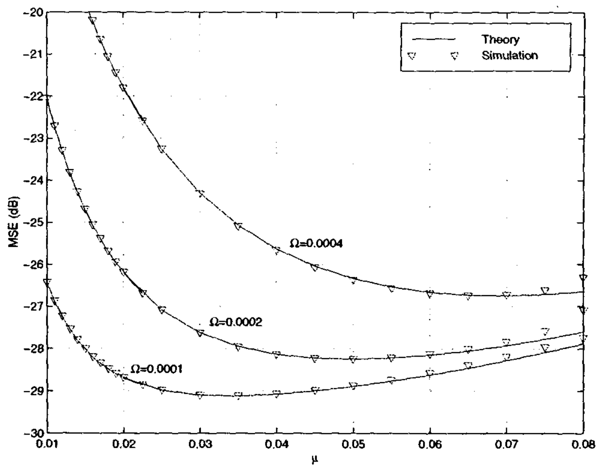

Figure 1: *Theoretical and experimental* EMSE *versus* μ.

Ω=0.0001, we can see that the experimental MSE possesses a well-defined minimum at μ=0.035, which is close to the estimate provided by (20) — μ_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 is significantly affected by the frequency offset Ω, even for very small values of Ω!

Acknowledgment. The authors would like to thank Dr. M. Rupp for bringing to their attention the work [1].

5. REFERENCES

[1] M. Rupp. LMS tracking behavior under periodically changing systems. *Proc. European Signal Processing Conference*, Island of Rhodes, Greece, Sep. 1998.

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

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

[4] 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.

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

[6] 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, Jun. 1999.

[7] N. R. Yousef and A. H. Sayed. Tracking analysis of the LMF and LMMN adaptive algorithms. *Proc. of the 33rd Asilomar Conference on Signals, Systems, and Computers*, Monterey, CA, Oct. 1999.

[8] R. Price. A useful theorem for nonlinear devices having Gaussian inputs. *IRE Trans. on Information Theory*, vol. 4, pp. 69–72, Jun. 1958.