PROC. 4TH IEEE-EURASIP WORKSHOP ON NONLINEAR SIGNAL AND IMAGE PROCESSING, VOL. 2, PP. 699–703, ANTALYA, TURKEY, JUNE 1999.

A UNIFIED APPROACH TO THE STEADY-STATE AND TRACKING ANALYSES OF ADAPTIVE FILTERING ALGORITHMS

NABIL R. YOUSEF AND ALI H. SAYED

Adaptive and Nonlinear Systems Laboratory
Electrical Engineering Department
University of California
Los Angeles, CA 90024

ABSTRACT

Most adaptive filters are inherently nonlinear and time variant systems. The nonlinearities in the update equations of these filters usually lead to significant difficulties in the study of their performance. This paper develops a new feedback approach to the steady-state and tracking analyses of adaptive algorithms that bypasses many of the difficulties encountered in traditional approaches. In this new formulation, we not only re-derive several earlier results in the literature, but we often do so under weaker assumptions, in a considerably more compact way, and we also obtain new results.

1. INTRODUCTION

This paper develops a new approach to the analysis of the steady-state performance of adaptive schemes. The approach is based on showing how a generic adaptive filter can be represented as a cascade of elementary sections, with each section consisting of a lossless system in the feedforward path and a feedback interconnection. By studying the energy flow through the cascade, we are able to establish a fundamental error variance relation. This relation has several ramifications, one of which is in the context of steady-state and tracking analyses, as we show in this paper.[1]

Thus consider noisy measurements $\{d(i)\}$ that arise from the linear model

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

where $\mathbf{w}^o$ is an unknown *column* vector that we wish to estimate, $v(i)$ accounts for measurement noise and modeling errors, and $\mathbf{u}_i$ denotes a nonzero *row* input (regressor) vector. Many adaptive algorithms have been developed in the literature for the estimation of $\mathbf{w}^o$ in different contexts (*e.g.*, echo cancelation, system identification, blind and non-blind channel equalization). In this paper, we focus on the following general class of algorithms:

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

where $\mathbf{w}_i$ is an estimate for $\mathbf{w}^o$ at iteration i, μ is the step-size, and $\mathbf{F}_e(i)$ denotes a generic scalar function of the quantities $\{\mathbf{u}_i, \mathbf{w}_i, d(i)\}$. Usually, $\mathbf{F}_e(i)$ is a (linear or nonlinear) function of the so-called output estimation error, defined by

$$e(i) = d(i) - \mathbf{u}_i \mathbf{w}_i \ . \tag{3}$$

Different choices for $\mathbf{F}_e(i)$ result in different adaptive algorithms. For example, Tab. 1 defines $\mathbf{F}_e(i)$ for many famous special cases of (2), for both blind and non-blind modes of adaptation. In the table, δ, R_1, and R_2 are constants, and $y(i) = \mathbf{u}_i \mathbf{w}_i$ is the adaptive filter output.

Table 1: *Examples for* $\mathbf{F}_e(i)$.

ALGORITHM	$\mathbf{F}_e(i)$
LMS	$e(i)$
NLMS	$e(i)/\Vert\mathbf{u}_i\Vert^2$
LMF	$e^3(i)$
LMMN	$\delta e(i) + (1-\delta)e^3(i)$
SA	$\text{sign}[e(i)]$
CMA1-2	$[R_1 \frac{y(i)}{\vert y(i)\vert} - y(i)]$
CMA2-2	$y(i)[R_2 - \vert y(i)\vert^2]$

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 $\tilde{\mathbf{w}}_i = \mathbf{w}^o - \mathbf{w}_i$ denotes the weight error vector. Under the often realistic assumption that (see, *e.g.*, [1]–[4]):

[1]This work was supported in part by the National Science Foundation under Award No. CCR-9732376.

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

$$\text{MSE} = \sigma_v^2 + \lim_{i\to\infty} \mathrm{E}\,|\mathbf{u}_i\tilde{\mathbf{w}}_i|^2 \quad . \tag{4}$$

Now the standard way for evaluating (4), and which dominates most derivations in the literature, is the following. First, one assumes, in addition to A.1, that the regression vector $\mathbf{u}_i$ is independent of $\tilde{\mathbf{w}}_i$. Then the above MSE becomes

$$\text{MSE} = \sigma_v^2 + \lim_{i\to\infty} \mathrm{Tr}(\mathbf{R}\mathbf{C}_i) \quad , \tag{5}$$

where $\mathbf{C}_i$ denotes the weight error covariance matrix, $\mathbf{C}_i = \mathrm{E}\,\tilde{\mathbf{w}}_i\tilde{\mathbf{w}}_i^*$, and $\mathbf{R} = \mathrm{E}\,\mathbf{u}_i^*\mathbf{u}_i$ is the input covariance matrix. As is evident from (5), this method of computation requires the determination of the steady-state value of $\mathbf{C}_i$, say $\mathbf{C}_\infty$. Finding $\mathbf{C}_\infty$ is a burden, especially for adaptive schemes with nonlinear update equations, which is the case for most of the algorithms listed in Tab. 1. This explains why the steady-state analysis of these algorithms in the literature is more advanced in some cases than in others. It also explains why such analyses have often been carried out separately for each individual algorithm and under varied assumptions. However, it would be very useful to develop a unifying framework that can handle a variety of algorithms. This paper takes an important step in this direction. More specifically, the following are the novel contributions of this work:

1. We develop a new feedback approach for evaluating the MSE of a large class of adaptive schemes. This approach distinguishes itself from earlier approaches in that it bypasses the need for working directly with $\mathbf{C}_i$ or with its limiting value.

2. The feedback approach not only allows us to re-derive several earlier results in literature in a unified manner, but it does so with considerably less effort and often under weaker assumptions.

3. The approach also allows us to derive several new results, especially for adaptive filters with nonlinear updates for which approaches that require $\mathbf{C}_i$ are not easily applicable.

4. The approach further establishes the significant conclusion that the tracking analysis of adaptive schemes can be obtained almost by inspection from the results in the stationary case. In contrast, analyses for both the stationary and non-stationary cases have always been carried out separately in the literature.

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,

$$e_a(i) = \mathbf{u}_i\tilde{\mathbf{w}}_i \ , \quad e_p(i) = \mathbf{u}_i\tilde{\mathbf{w}}_{i+1} \ .$$

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 $\mathbf{w}^o$ 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:

$$e_p(i) = e_a(i) - \frac{\mu}{\bar{\mu}(i)}\mathbf{F}_e(i) \ , \tag{6}$$

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

$$\tilde{\mathbf{w}}_{i+1} = \tilde{\mathbf{w}}_i - \bar{\mu}(i)\mathbf{u}_i^*[e_a(i) - e_p(i)] \ . \tag{7}$$

By evaluating the energies of both sides of this equation we obtain [5, 6]

$$\|\tilde{\mathbf{w}}_{i+1}\|^2 + \bar{\mu}(i)|e_a(i)|^2 = \|\tilde{\mathbf{w}}_i\|^2 + \bar{\mu}(i)|e_p(i)|^2 \ . \tag{8}$$

This energy conservation relation holds for all adaptive algorithms whose recursions are of the form given by (2). *No approximations or assumptions are needed to establish (8)*; it is an exact relation that 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. The relation also has an interesting system-theoretic interpretation. It establishes that the mapping from $\left\{\tilde{\mathbf{w}}_i, \sqrt{\bar{\mu}(i)}e_p(i)\right\}$ to $\left\{\tilde{\mathbf{w}}_{i+1}, \sqrt{\bar{\mu}(i)}e_a(i)\right\}$ is energy preserving (or lossless). Furthermore, combining (8) with (6), we see that both relations establish the existence of the feedback configuration shown in Fig. 1, where $\mathcal{T}$ denotes a lossless map and q^{-1} denotes the unit delay operator. [The variable $\mathbf{q}_i$ that appears in the figure should be set to zero at this first part of the paper. It will be nonzero when we discuss later the tracking performance of an adaptive filter. We use the same figure for both cases to emphasize that they will only differ by one additional disturbance represented by $\mathbf{q}_i$.]

Relevance to Steady-State Performance Analysis

As mentioned in the introduction, relation (8) has several ramifications. It was used in [5, 6] (and in some of the references therein) to study the robustness and l_2-stability of adaptive filters. Here we show its significance to steady-state analysis.

Recall that we are interested in evaluating the MSE of an adaptive filter once it reaches steady-state. To do so, we simply note that $\mathrm{E}\,\|\tilde{\mathbf{w}}_{i+1}\|^2 = \mathrm{E}\,\|\tilde{\mathbf{w}}_i\|^2$ in steady-state, so

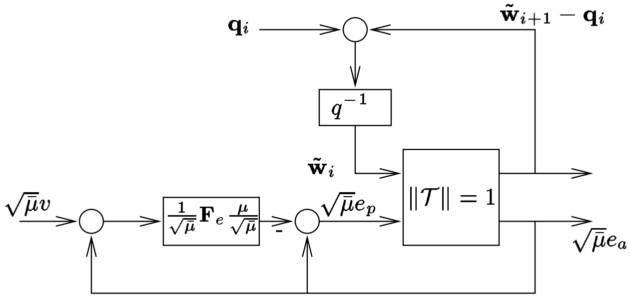

Figure 1: *Lossless mapping and a feedback loop.*

that by taking expectations of both sides of (8) we obtain the equality

$$\mathrm{E}\,\bar{\mu}(i)|e_a(i)|^2 = \mathrm{E}\,\bar{\mu}(i)|e_p(i)|^2 \ .$$

Using (6), the above collapses to the following *fundamental error variance relation* in terms of $\{e_a(i), v(i)\}$ only (recall that $e(i) = e_a(i) + v(i)$):

$$\mathrm{E}\,\bar{\mu}(i)|e_a(i)|^2 = \mathrm{E}\,\bar{\mu}(i)\left|e_a(i) - \frac{\mu}{\bar{\mu}(i)}\mathbf{F}_e(i)\right|^2 \ . \tag{9}$$

This equation can now be solved for the steady-state excess mean-square-error (EMSE):

$$\zeta \stackrel{\Delta}{=} \lim_{i\to\infty} \mathrm{E}\,|e_a(i)|^2 \ .$$

Observe from (4) that the desired MSE is given by $\mathsf{MSE} = \sigma_v^2 + \zeta$, so that finding ζ is equivalent to finding the MSE. Moreover, observe again that (9) is an *exact* relation that holds without any approximations or assumptions (except for the assumption that the filter is in steady-state).

The procedure of finding the EMSE through (9) completely avoids the need for evaluating $\mathrm{E}\,\|\tilde{\mathbf{w}}_\infty\|^2$. This is because in steady-state, and in view of the energy-preserving relation (8), the effect of the weight error variance is canceled out!

3. STEADY-STATE ANALYSIS

We now apply the above general procedure to various adaptive algorithms from Tab. 1. 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 (9).

3.1. The LMS Algorithm

For LMS we have $\mathbf{F}_e(i) = e(i) = e_a(i) + v(i)$. Substituting into (9) and using A.1, it follows immediately that

$$2\mu\zeta^{\mathsf{LMS}} = \mu^2\,\mathrm{E}\left(\|\mathbf{u}_i\|^2\,|e_a(i)|^2\right) + \mu^2\sigma_v^2\,\mathrm{Tr}(\mathbf{R}) \ . \tag{10}$$

To solve for ζ^{LMS} we consider two 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 relative to the second term on the right-hand side of (10), so that

$$\zeta^{\mathsf{LMS}} = \frac{\mu}{2}\sigma_v^2\,\mathrm{Tr}(\mathbf{R}). \tag{11}$$

This is the same result obtained in [7] for small values of μ but here we get it in a much simpler way.

2. For larger values of μ, equation (10) can be solved by imposing the following (often studied) assumption:

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

This assumption in fact becomes very realistic for long filter lengths. Furthermore, it becomes *exact* for the case of constant modulus data that arises in many adaptive filtering applications. Using A.2, and (10), we directly obtain

$$\zeta^{\mathsf{LMS}} = \frac{\mu\sigma_v^2\,\mathrm{Tr}(\mathbf{R})}{2 - \mu\,\mathrm{Tr}(\mathbf{R})} \ . \tag{12}$$

This is also a well-known result (see [8]) but is obtained here very directly and under the single assumption A.2.

3.2. The LMF and LMMN Algorithms

For the least-mean mixed-norm (LMMN) algorithm, we have $\mathbf{F}_e(i) = \delta e(i) + (1-\delta)e^3(i)$. The least-mean fourth (LMF) algorithm corresponds to the special case $\delta = 0$. Introduce, for compactness of notation,

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

We again consider two cases. In both cases, we make the reasonable assumption that in steady-state $|e_a(i)|^2 \ll |v(i)|^2$. We also assume A.1.

1. For values of μ that are small enough so that the term $\mu^2\,\mathrm{E}\,\|\mathbf{u}_i\|^2|e_a(i)|^2$ could be ignored, we obtain

$$\zeta^{\mathsf{LMMN}} = \frac{\mu}{2}\left(\frac{\delta^2\sigma_v^2 + 2\delta\bar{\delta}\xi_v^4 + \bar{\delta}^2\xi_v^6}{\delta + 3\bar{\delta}\sigma_v^2}\right)\mathrm{Tr}(\mathbf{R}) \ . \tag{13}$$

This is the same result obtained in [9], but we get it here more directly and under weaker conditions. For $\delta = 0$, the above expression collapses to

$$\zeta^{\mathsf{LMF}} = \frac{\mu}{2}\left(\frac{\xi_v^6}{3\sigma_v^2}\right)\mathrm{Tr}(\mathbf{R}) \ , \tag{14}$$

which is the same expression obtained in [10] by using the so-called independence assumptions.

2. For larger values of μ, and using A.2 again, we get the following new expressions for the EMSE:

$$\zeta^{\mathsf{LMMN}} = \frac{\mu(\delta^2\sigma_v^2 + 2\delta\bar{\delta}\xi_v^4 + \bar{\delta}^2\xi_v^6)\operatorname{Tr}(\mathbf{R})}{2(\delta + 3\bar{\delta}\sigma_v^2) - \mu(\delta^2 + 6\delta\bar{\delta}\sigma_v^2 + 9\bar{\delta}\xi_v^4)\operatorname{Tr}(\mathbf{R})}$$

$$\zeta^{\mathsf{LMF}} = \frac{\mu\xi_v^6\operatorname{Tr}(\mathbf{R})}{6\sigma_v^2 - 9\mu\xi_v^4\operatorname{Tr}(\mathbf{R})} \ .$$

3.3. The NLMS Algorithm

For the normalized LMS algorithm, $\mathbf{F}_e(i) = e(i)/\|\mathbf{u}_i\|^2$. In this case, relation (9), and assumption A.1, lead to the equality

$$(2-\mu)\operatorname{E}\left(\frac{|e_a(i)|^2}{\|\mathbf{u}_i\|^2}\right) = \mu\sigma_v^2\operatorname{E}\left(\frac{1}{\|\mathbf{u}_i\|^2}\right) \ . \qquad (15)$$

Again this is an exact equality. We consider two cases.

1. Under assumption A.2, we have

$$\operatorname{E}\left(\frac{|e_a(i)|^2}{\mu^2\|\mathbf{u}_i\|^2}\right) \ = \ \operatorname{E}|e_a(i)|^2 \ \operatorname{E}\left(\frac{1}{\mu^2\|\mathbf{u}_i\|^2}\right) \ ,$$

so that (15) leads to the expression

$$\zeta^{\mathsf{NLMS}} = \frac{\mu\sigma_v^2}{(2-\mu)} \ . \qquad (16)$$

This result is in fact *exact* for constant modulus data. Observe also that it is independent of $\mathbf{R}$.

2. In some works (see, *e.g.*, p. 443 of [2]), the following approximation is instead used

$$\operatorname{E}\left(\frac{|e_a(i)|^2}{\|\mathbf{u}_i\|^2}\right) \ \approx \frac{\operatorname{E}|e_a(i)|^2}{\operatorname{E}\|\mathbf{u}_i\|^2} \ ,$$

in which case (15) leads to

$$\zeta^{\mathsf{NLMS}} = \frac{\mu\sigma_v^2}{(2-\mu)} \ \operatorname{E}\left(\frac{1}{\|\mathbf{u}_i\|^2}\right)\operatorname{Tr}(\mathbf{R}) \ . \qquad (17)$$

This is the same expression obtained in [13] in a very different (and also less direct) way.

3.4. The Sign Algorithm

For the sign algorithm (SA), we have $\mathbf{F}_e(i) = \operatorname{sign}[e(i)]$. In this case, relation (9) leads to the equality:

$$\operatorname{E}[e_a(i)\operatorname{sign}(e_a(i) + v(i))] = \frac{\mu}{2}\operatorname{Tr}(\mathbf{R}) \ . \qquad (18)$$

By assuming that $e_a(i)$ and $v(i)$ are jointly Gaussian, and by using A.1 and Price's theorem [11], we obtain

$$\operatorname{E}[e_a(i)\operatorname{sign}(e_a(i) + v(i))] = \sqrt{\frac{2}{\pi}}\frac{\operatorname{E}|e_a(i)|^2}{\sqrt{\sigma_v^2 + \operatorname{E}|e_a(i)|^2}} \ .$$

Substituting into (18) and solving for $\operatorname{E}|e_a(i)|^2$, we find that

$$\zeta^{\mathsf{SA}} = \frac{\alpha}{2}\cdot\left(\alpha + \sqrt{\alpha^2 + 4\sigma_v^2}\right) \ , \qquad (19)$$

where $\alpha = \sqrt{\frac{\pi}{8}}\mu\operatorname{Tr}(\mathbf{R})$. This is the same result that was obtained in [12] by using the independence assumptions. Here we have shown that the same result holds without any independence assumptions!

3.5. The CM Algorithms

Similar analyses can be carried out for constant modulus (CM) algorithms. The details are provided in [14]. Here we only briefly comment on one particular case for the sake of illustration. Assume $v(i) = 0$ (and, hence, $e_a(i) = e(i)$) and define

$$\sigma_d^2 = \operatorname{E}|d(i)|^2 \ , \ \ \xi_d^4 = \operatorname{E}|d(i)|^4 \ , \ \ \xi_d^6 = \operatorname{E}|d(i)|^6 \ .$$

Let $R_2 = \xi_d^4/\sigma_d^2$ and assume also that all data are real-valued (the complex case is studied in [14]). Define further, for compactness of notation, $z(i) = y(i)(R_2 - |y(i)|^2)$. Then relation (9) yields for CMA2-2,

$$2\mu\operatorname{E}e_a(i)z(i) = \mu^2\operatorname{E}[\|\mathbf{u}_i\|^2|z(i)|^2] \ .$$

To solve this equation for $\operatorname{E}|e_a(i)|^2$, we make the following reasonable (and common) assumption.

<u>A.3</u> The signals $d(i)$ and $e_a(i)$ are independent in steady-state so that $\operatorname{E}d(i)e_a(i) = 0$, since the signal $d(i)$ is assumed zero mean.

Using assumptions A.2 and A.3, yields for small enough μ:

$$\zeta^{\mathsf{CMA2-2}} \ = \ \frac{\mu}{2}\left(\frac{\sigma_d^2R_2^2 - 2R_2\xi_d^4 + \xi_d^6}{2(3\sigma_d^2 - R_2)}\right)\operatorname{Tr}(\mathbf{R}) \ \ .$$

This is a slightly different expression from the one obtained in [15] via a different (and less direct) route. It was shown in [14] that the above expression leads to a better approximation for the MSE. More discussion on, and new EMSE expressions for, other CM algorithms can be found in [14].

4. TRACKING ANALYSIS

In a nonstationary environment, the data $\{d(i)\}$ is assumed to arise from a linear model of the form $d(i) = \mathbf{u}_i\mathbf{w}_i^o + v(i)$, where the unknown system $\mathbf{w}_i^o$ is now time-variant. It is often assumed that the variation in $\mathbf{w}_i^o$ is according to the model $\mathbf{w}_{i+1}^o = \mathbf{w}_i^o + \mathbf{q}_i$, where $\mathbf{q}_i$ denotes the random perturbation. The purpose of the tracking analysis of an adaptive filter is to study its ability to track such time-variations. We now show how to evaluate the tracking performance of

Table 2: *Expressions for the EMSE in a nonstationary environment and small μ.*

ALGORITHM	EMSE	ASSUMPTIONS
LMS	$\frac{1}{2}\mu^{-1}\,\mathrm{Tr}(\mathbf{Q}) + \frac{1}{2}\mu\sigma_v^2\,\mathrm{Tr}(\mathbf{R})$	1, 4, 5.
NLMS	$\frac{\mu^{-1}}{2-\mu}\,\mathrm{Tr}(\mathbf{R})\,\mathrm{Tr}(\mathbf{Q}) \;+\; \frac{\mu\sigma_v^2}{2-\mu}\,\mathrm{E}\left(\frac{1}{\Vert\mathbf{u}_i\Vert^2}\right)\mathrm{Tr}(\mathbf{R})$	1, 2, 4,5
LMF	$\frac{\mu^{-1}\,\mathrm{Tr}(\mathbf{Q})+\mu\xi_v^6\,\mathrm{Tr}(\mathbf{R})}{6\sigma_v^2}$	1, 4, 5
LMMN	$\frac{\mu^{-1}\,\mathrm{Tr}(\mathbf{Q})+\mu[\delta^2\sigma_v^2+2\delta\bar{\delta}\xi_v^4+\bar{\delta}\xi_v^6]\,\mathrm{Tr}(\mathbf{R})}{2[\delta+3\bar{\delta}\sigma_v^2]}$	1, 4, 5
SA	$\frac{\beta}{2}\left(\beta+\sqrt{\beta^2+4\sigma_v^2}\right),\ \beta=\sqrt{\frac{\pi}{8}}\left[\mu^{-1}\,\mathrm{Tr}(\mathbf{Q})+\mu\,\mathrm{Tr}(\mathbf{R})\right]$	1, 4, 5, Gaussian errors
CMA2-2	$\frac{\mu^{-1}\,\mathrm{Tr}(\mathbf{Q})+\mu\left(\sigma_d^2R_2^2-2R_2\xi_d^4+\xi_d^6\right)\mathrm{Tr}(\mathbf{R})}{2(3\sigma_d^2-R_2)}$	1, 2, 3, 4, 5

an adaptive algorithm by the same feedback method proposed in this paper.

For this purpose, we first redefine the weight error vector as $\tilde{\mathbf{w}}_i = \mathbf{w}_i^o - \mathbf{w}_i$, and the a-posteriori estimation error as $e_p(i) = \mathbf{u}_i\left(\tilde{\mathbf{w}}_{i+1} - \mathbf{q}_i\right)$. Then $\tilde{\mathbf{w}}_i$ satisfies

$$\tilde{\mathbf{w}}_{i+1} = \tilde{\mathbf{w}}_i - \mu(i)\ \mathbf{u}_i^*\ \mathbf{F}_e(i) + \mathbf{q}_i\ . \tag{20}$$

If we further multiply (20) by $\mathbf{u}_i$ from the left, we obtain that (6) and (7) still hold for the nonstationary case, while (8) becomes:

$$\Vert\tilde{\mathbf{w}}_{i+1} - \mathbf{q}_i\Vert^2 + \bar{\mu}(i)|e_a(i)|^2 = \Vert\tilde{\mathbf{w}}_i\Vert^2 + \bar{\mu}(i)|e_p(i)|^2\ . \tag{21}$$

For mathematical tractability of the tracking analysis, we impose the following assumptions, which are typical in the context of tracking analysis of adaptive filters (see, *e.g.*, [8]).

<u>A.4</u> The sequences $\{\mathbf{u}_i\}$ and $\{v(i)\}$ are mutually statistically independent of $\{\mathbf{q}_i\}$.

<u>A.5</u> 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 (6), A.4 and A.5, it is straightforward to verify that the variance relation (9) should now be replaced by:

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

Comparing the above with (9), we see that evaluating the nonstationary EMSE is simply a straightforward extension of evaluating the stationary EMSE! The only addition is the steady-state contribution by the system nonstationarity, which is equal to $\mathrm{Tr}(\mathbf{Q})$.

This is a very significant and helpful observation in the context of the tracking analysis of adaptive algorithms, since it allows us to arrive at tracking results almost by inspection from the stationary case results. In the literature, both cases have always been studied separately. We summarize the EMSE results for tracking in Tab. 2 for the case of small μ (for brevity). The expressions for LMF, LMMN, and CMA2-2 are new.

Moreover, by differentiating the EMSE expressions in Tab. 2 with respect to μ, we obtain several new expressions for the optimum step-sizes that achieve the lowest EMSE. Due to space limitations, we do not list these expressions here.

5. REFERENCES

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

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

[3] V. Solo and X. Kong. *Adaptive Signal Processing Algorithms*. Englewood Cliffs, NJ, 1995.

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

[5] 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, June 1996.

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

[7] S. K. Jones, R. K. Cavin, and W. M. Reed. Analysis of error-gradient adaptive linear estimators for a class of stationary dependent processes. *IEEE Transactions on Information Theory*, vol. 28, pp. 318–329, March 1982.

[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] 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.

[10] 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.

[11] R. Price. A useful theorem for nonlinear devices having Gaussian inputs. *IRE Trans. on Inform. Theory*, vol. IT-4, pp. 69–72, June 1958.

[12] V. Mathews, and S. Cho. Improved convergence analysis of stochastic gradient adaptive filters using the sign algorithm. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 35, no. 4, April 1987.

[13] D. T. M. Slock. On the convergence behavior of the LMS and the normalized LMS algorithms. *IEEE Transactions on Signal Processing*, vol. 41, no. 9, pp. 2811–2825, Sep. 1993.

[14] J. Mai and A. H. Sayed. A feedback approach to the steady-state performance of fractionally-spaced blind adaptive equalizers. *Submitted for publication*.

[15] I. Fijakow, C. E. Manlove, and C. R. Johnson. Adaptive fractionally spaced blind CMA equalization: Excess MSE. *IEEE Transactions on Signal Processing*, vol. 46, no. 1, pp. 227–231, January 1998.