TRANSIENT ANALYSIS OF ADAPTIVE FILTERS

TAREQ Y. AL-NAFFOURI[1] AND ALI H. SAYED[2]

[1]Electrical Engineering Department
Stanford University, CA 94305

[2]Electrical Engineering Department
University of California, Los Angeles, CA 90095

ABSTRACT

This paper develops a framework for the mean-square analysis of adaptive filters with general data and error nonlinearities. The approach relies on energy conservation arguments and is carried out without restrictions on the probability distribution of the input sequence. In particular, for adaptive filters with diagonal matrix nonlinearities, we provide closed form expressions for the steady-state performance and necessary and sufficient conditions for stability. We carry out a similar study for long adaptive filters that employ error nonlinearities relying on a weaker form of the independence assumption. We provide expressions for the steady-state error and bounds on the step-size for stability by exploiting the Cramer-Rao bound of the underlying estimation process.

1. ADAPTIVE FILTERING MODEL

Consider noisy measurements $d(i) = \mathbf{u}_i \boldsymbol{w}^o + v(i)$, where $\boldsymbol{w}^o$ denotes an unknown column vector that we wish to estimate, $\mathbf{u}_i$ is a row regression vector, and $v(i)$ is measurement noise. Adaptive schemes for estimating $\boldsymbol{w}^o$ rely on recursive updates of the general form

$$\boxed{\boldsymbol{w}_{i+1} = \boldsymbol{w}_i + \mu \boldsymbol{H}(\mathbf{u}_i)\mathbf{u}_i^T f(e(i)), \quad i \geq 0} \tag{1}$$

where $\boldsymbol{w}_i$ is the estimate of $\boldsymbol{w}^o$ at time i, μ is the step-size, and

$$\boxed{e(i) = d(i) - \mathbf{u}_i \boldsymbol{w}_i} \tag{2}$$

is the estimation error. The correction term in (1) is usually expressed in a separable form, $\boldsymbol{H}(\mathbf{u}_i)\mathbf{u}_i^T f(e(i))$,

This work was partially supported by the National Science Foundation under awards ECS-9820765 and CCR-9732376. The work of T. Y. Al-Naffouri was also partially supported by a fellowship from King Fahd University of Petroleum and Minerals, Saudi Arabia.

Table 1: Examples for $f(e(i))$ and $\boldsymbol{H}(\mathbf{u}_i)$

ALGORITHM	$f[e(i)]$
LMS	$e(i)$
LMF	$e^3(i)$
LMF family	$e^{2k+1}(i)$
LMMN	$ae(i) + be^3(i)$
Sat. nonlin.	$\int_0^{e(i)} \exp\left(-\frac{z^2}{2\sigma_z^2}\right) dz$
Sign error	$\text{sign}[e(i)]$

ALGORITHM	$\boldsymbol{H}(\mathbf{u}_i)$
NLMS	$\frac{1}{\lVert\mathbf{u}_i\rVert^2}\boldsymbol{I}$
ϵ-NLMS	$\frac{1}{\epsilon+\lVert\mathbf{u}_i\rVert^2}\boldsymbol{I}$
sign regressor	$\text{diag}\left(\frac{\text{sign}(u_{i_1})}{u_{i_1}}, \ldots, \frac{\text{sign}(u_{i_M})}{u_{i_M}}\right)$
variable steps	$\text{diag}(\mu_1, \mu_2, \ldots, \mu_M)$

where $f(e(i))$ denotes a scalar error nonlinearity and $\boldsymbol{H}(\mathbf{u}_i)$ denotes a data nonlinearity and is taken as a diagonal matrix with nonnegative entries. In this paper, we focus on correction terms that are nonlinear in the data or in the error but not both. This class of algorithms is general enough to include the special cases listed in Table 1. Several of these algorithms were already considered in the literature (see, e.g., [1]–[3] and [6] and the many references therein). The purpose of this article is to provide a framework for performing mean-square analysis of the general class of algorithms (1)–(2) in a unified manner. This is achieved by relying on the energy-conservation approach developed in [4]–[6] and by expanding it to handle both transient analysis and mean-square analysis.

2. ENERGY RELATION

Mean-square analysis of (1)-(2) is best carried out in terms of the normalized regressor $\overline{\mathbf{u}}_i = \mathbf{u}_i \boldsymbol{H}(\mathbf{u}_i)$ and the following error quantities:

$\tilde{\boldsymbol{w}}_i \triangleq \boldsymbol{w}^o - \boldsymbol{w}_i$	weight-error vector
$e_a^{\Sigma}(i) \triangleq \mathbf{u}_i \boldsymbol{\Sigma} \tilde{\boldsymbol{w}}_i$	weighted a priori error
$e_p^{\Sigma}(i) \triangleq \mathbf{u}_i \boldsymbol{\Sigma} \tilde{\boldsymbol{w}}_{i+1}$	weighted a posteriori error

where $\boldsymbol{\Sigma}$ denotes a weighting matrix. We reserve special notation for the case $\boldsymbol{\Sigma} = \boldsymbol{I}$: $e_a(i) = e_a^I(i)$ and $e_p(i) = e_p^I(i)$. Using these error quantities, we can rewrite the adaptive algorithm (1)-(2) as

$$\tilde{\boldsymbol{w}}_{i+1} = \tilde{\boldsymbol{w}}_i - \mu \overline{\mathbf{u}}_i^T f(e(i)) \tag{3}$$

$$e(i) = e_a(i) + v(i) \tag{4}$$

We also find it useful to use the compact notation $\|\tilde{\boldsymbol{w}}_i\|_{\Sigma}^2 = \tilde{\boldsymbol{w}}_i^T \boldsymbol{\Sigma} \tilde{\boldsymbol{w}}_i$. This notation is convenient because it enables us to transform operations on $\tilde{\boldsymbol{w}}_i$ into operations on the norm subscript, as demonstrated by the following properties. Let a_1 and a_2 be scalars and $\boldsymbol{\Sigma}_1$ and $\boldsymbol{\Sigma}_2$ be symmetric matrices of size M. Then

1) Superposition.

$$a_1 \|\tilde{\boldsymbol{w}}_i\|_{\Sigma_1}^2 + a_2 \|\tilde{\boldsymbol{w}}_i\|_{\Sigma_2}^2 = \|\tilde{\boldsymbol{w}}_i\|_{a_1 \Sigma_1 + a_2 \Sigma_2}^2$$

2) Polarization.

$$(\mathbf{u}_i \boldsymbol{\Sigma}_1 \tilde{\boldsymbol{w}}_i)(\mathbf{u}_i \boldsymbol{\Sigma}_2 \tilde{\boldsymbol{w}}_i) = \|\tilde{\boldsymbol{w}}_i\|_{\Sigma_1 u_i^T u_i \Sigma_2}^2$$

3) Independence. If $\tilde{\boldsymbol{w}}_i$ and $\mathbf{u}_i$ are independent,

$$E\left[\|\tilde{\boldsymbol{w}}_i\|_{\Sigma_1 u_i^T u_i \Sigma_2}^2\right] = E\left[\|\tilde{\boldsymbol{w}}_i\|_{\Sigma_1 E[u_i^T u_i] \Sigma_2}^2\right]$$

4) Notational convention. Using the vector notation, we shall write $\|\tilde{\boldsymbol{w}}_i\|_{\text{vec}(\Sigma_1)}^2 \triangleq \|\tilde{\boldsymbol{w}}_i\|_{\Sigma_1}^2$

With the above definitions and notation at hand, we proceed to premultiply both sides of (3) by $\mathbf{u}_i \boldsymbol{H}(\mathbf{u}_i) \boldsymbol{\Sigma}$ to get

$$\mathbf{u}_i \boldsymbol{H}(\mathbf{u}_i) \boldsymbol{\Sigma} \tilde{\boldsymbol{w}}_{i+1} = \mathbf{u}_i \boldsymbol{H}(\mathbf{u}_i) \boldsymbol{\Sigma} \tilde{\boldsymbol{w}}_i - \mu f(e(i)) \mathbf{u}_i \boldsymbol{H}(\mathbf{u}_i) \boldsymbol{\Sigma} \overline{\mathbf{u}}_i^T$$

Incorporating the expressions for $\overline{\mathbf{u}}_i$, $e_a^{(\cdot)}$, and $e_p^{(\cdot)}$, and solving for $\mu f(e(i))$, we find that

$$\mu f(e(i)) = \frac{e_a^{H\Sigma}(i)}{\|\overline{\mathbf{u}}_i\|_{\Sigma}^2} - \frac{e_p^{H\Sigma}(i)}{\|\overline{\mathbf{u}}_i\|_{\Sigma}^2} \tag{5}$$

Combining (3) and (5) to eliminate $\mu f(e(i))$, and taking the $\boldsymbol{\Sigma}$-weight of the resulting expression leads to the energy conservation relation:

$$\boxed{\|\tilde{\boldsymbol{w}}_{i+1}\|_{\Sigma}^2 + \frac{\left|e_a^{H\Sigma}(i)\right|^2}{\|\overline{\mathbf{u}}_i\|_{\Sigma}^2} = \|\tilde{\boldsymbol{w}}_i\|_{\Sigma}^2 + \frac{\left|e_p^{H\Sigma}(i)\right|^2}{\|\overline{\mathbf{u}}_i\|_{\Sigma}^2}} \tag{6}$$

This equality relates the weighted energies of the error variables $\{\tilde{\boldsymbol{w}}_i, \tilde{\boldsymbol{w}}_{i+1}, e_a^{\Sigma}(i), e_p^{\Sigma}(i)\}$; it is the weighted version of the energy relation derived in [4]–[6] and used there, and in other related references, to study the performance of adaptive filters from both deterministic and stochastic points of view. The inclusion of the weighting factor Σ allows us to perform both transient and steady-state analyses. Observe that no assumptions or approximations were used to derive (6). This relation will be the starting point for much of the subsequent discussion.

3. THE DATA NONLINEARITY CASE

In this section, we assume $f(e(i)) = e(i)$ and proceed to study the mean-square performance of the resulting algorithm. For this purpose, we rely on the following independence assumptions:

AN The noise $v(i)$ is i.i.d. and independent of the input.

AI The sequence of regressors $\{\mathbf{u}_i\}$ is independent with zero mean and autocorrelation matrix $\mathbf{R}$.

Thus note first that (5) becomes

$$e_p^{H\Sigma}(i) = e_a^{H\Sigma}(i) - \mu e(i) \|\overline{\mathbf{u}}_i\|_{\Sigma}^2$$

Substituting this expression for $e_p^{H\Sigma}(i)$ into the energy relation (6), we get

$$\|\tilde{\boldsymbol{w}}_{i+1}\|_{\Sigma}^2 = \|\tilde{\boldsymbol{w}}_i\|_{\Sigma}^2 - 2\mu e_a^{H\Sigma}(i) e(i) + \mu^2 \|\overline{\mathbf{u}}_i\|_{\Sigma}^2 e^2(i) \tag{7}$$

By further incorporating (4) and assumption AN, (7) reads under expectation

$$E\left[\|\tilde{\boldsymbol{w}}_{i+1}\|_{\Sigma}^2\right] = E\left[\|\tilde{\boldsymbol{w}}_i\|_{\Sigma}^2\right] - 2\mu E\left[e_a^{H\Sigma}(i) e_a(i)\right] + \mu^2 E\left[e_a^2(i) \|\overline{\mathbf{u}}_i\|_{\Sigma}^2\right] + \mu^2 \sigma_v^2 E\left[\|\overline{\mathbf{u}}_i\|_{\Sigma}^2\right] \tag{8}$$

Using the weighted-norm properties, we can rewrite the estimation error expectations in (8) as some weighted norms of $\tilde{\boldsymbol{w}}_i$:

$$2 e_a(i) e_a^{H\Sigma}(i) = 2 \tilde{\boldsymbol{w}}_i^T \mathbf{u}_i^T \overline{\mathbf{u}}_i \boldsymbol{\Sigma} \tilde{\boldsymbol{w}}_i = \|\tilde{\boldsymbol{w}}_i\|_{u_i^T \overline{u}_i \Sigma + \Sigma \overline{u}_i^T u_i}^2 \tag{9}$$

$$e_a^2 \|\overline{\mathbf{u}}_i\|_{\Sigma}^2 = \tilde{\boldsymbol{w}}_i^T \mathbf{u}_i^T \|\overline{\mathbf{u}}_i\|_{\Sigma}^2 \mathbf{u}_i \tilde{\boldsymbol{w}}_i = \|\tilde{\boldsymbol{w}}_i\|_{u_i^T \|\overline{u}_i\|_{\Sigma}^2 u_i}^2 \tag{10}$$

Substituting (9)–(10) into (8) and using assumption AI yields

$$E\left[\|\tilde{\boldsymbol{w}}_{i+1}\|_{\Sigma}^2\right] = E\left[\|\tilde{\boldsymbol{w}}_i\|_{\Sigma}^2\right] + \mu^2 E\left[\|\tilde{\boldsymbol{w}}_i\|_{E\left[u_i^T \|\overline{u}_i\|_{\Sigma}^2 u_i\right]}^2\right]$$

$$-\mu E\left[\|\tilde{\boldsymbol{w}}_i\|^2_{E[u_i\overline{u}_i]\Sigma+\Sigma E[\overline{u}_i u_i]}\right]+\mu^2\sigma_v^2 E\left[\|\overline{\mathbf{u}}_i\|^2_{\Sigma}\right]$$

or, more compactly,

$$E\left[\|\tilde{\boldsymbol{w}}_{i+1}\|^2_{\Sigma_{i+1}}\right]=E\left[\|\tilde{\boldsymbol{w}}_i\|^2_{\Sigma_i}\right]+\mu^2\sigma_v^2 E\left[\|\overline{\mathbf{u}}_i\|^2_{\Sigma_{i+1}}\right] \quad (11)$$

where a time index $(i+1)$ has been attached to $\boldsymbol{\Sigma}$, and where $\{\boldsymbol{\Sigma}_i,\boldsymbol{\Sigma}_{i+1}\}$ are related via

$$\begin{aligned}\boldsymbol{\Sigma}_i=\boldsymbol{\Sigma}_{i+1}-\mu\boldsymbol{\Sigma}_{i+1}E\left[\overline{\mathbf{u}}_i^T\mathbf{u}_i\right]-\mu E\left[\mathbf{u}_i^T\overline{\mathbf{u}}_i\right]\boldsymbol{\Sigma}_{i+1}\\ +\mu^2 E[\|\overline{\mathbf{u}}_i\|^2_{\Sigma_{i+1}}\mathbf{u}_i^T\mathbf{u}_i] \quad (12)\end{aligned}$$

Relations (11)–(12) (or, equivalently, (14)–(15) below) are the equivalent representations of the energy relation (6) under assumptions AN and AI. They can be used to derive conditions for mean-square stability, as well as expressions for the steady-state mean-square error and mean-square deviation of an adaptive filter. To see this, we start by noting that the recursion for $\boldsymbol{\Sigma}_i$ can be rewritten more compactly, using the vec operation and the Kronecker product notation, as

$$\boxed{\boldsymbol{\sigma}_i=\boldsymbol{F}\boldsymbol{\sigma}_{i+1}} \quad (13)$$

where

$$\boxed{\boldsymbol{F}=E\left[\left(\boldsymbol{I}-\mu\mathbf{u}_i^T\overline{\mathbf{u}}_i\right)\otimes\left(\boldsymbol{I}-\mu\mathbf{u}_i^T\overline{\mathbf{u}}_i\right)\right]} \quad (14)$$

and $\boldsymbol{\sigma}_i=\text{vec}\,(\boldsymbol{\Sigma}_i)$. In light of (13), relation (11) becomes

$$\boxed{E\left[\|\tilde{\boldsymbol{w}}_{i+1}\|^2_{\sigma_{i+1}}\right]=E\left[\|\tilde{\boldsymbol{w}}_i\|^2_{F\sigma_{i+1}}\right]+\mu^2\sigma_v^2 E\left[\|\overline{\mathbf{u}}_i\|^2_{\sigma_{i+1}}\right]} \quad (15)$$

By inspecting (15), it becomes clear that the recursion is stable if, and only if, the matrix $\boldsymbol{F}$ is stable. Thus let $\mathbf{A}=\boldsymbol{I}\otimes E[\mathbf{u}_i\overline{\mathbf{u}}_i]+E[\mathbf{u}_i\overline{\mathbf{u}}_i]\otimes\boldsymbol{I}$ and $B=E[\mathbf{u}_i\overline{\mathbf{u}}_i\otimes\mathbf{u}_i\overline{\mathbf{u}}_i]$. Then, from (14), $\boldsymbol{F}=\boldsymbol{I}-\mu\mathbf{A}+\mu^2\mathbf{B}$ and $\boldsymbol{F}$ will be stable if, and only if,

$$\boxed{0<\mu<\frac{1}{\lambda_{\max}(\mathbf{A}^{-1}\mathbf{B})}}$$

which provides the desired condition for mean-square stability.

Now assuming the filter is stable, we have

$$\lim_{i\to\infty}E\left[\|\tilde{\boldsymbol{w}}_{i+1}\|^2_{\sigma}\right]=\lim_{i\to\infty}E\left[\|\tilde{\boldsymbol{w}}_i\|^2_{\sigma}\right]$$

Thus, in the limit, and using the change of variables $\boldsymbol{\sigma}'=(\boldsymbol{I}-\boldsymbol{F})\boldsymbol{\sigma}$, relation (15) takes the form

$$\boxed{\lim_{i\to\infty}E\left[\|\tilde{\boldsymbol{w}}_i\|^2_{\sigma'}\right]=\mu^2\sigma_v^2 E\left[\|\overline{\mathbf{u}}_i\|^2_{(I-F)^{-1}\sigma'}\right]} \quad (16)$$

This expression allows us to evaluate the steady-state weight-error energy for any weight $\boldsymbol{\sigma}'$. In particular, we can get the mean-square error by choosing $\boldsymbol{\sigma}'=\text{vec}\,(\boldsymbol{R})$, and the mean-square deviation by choosing $\boldsymbol{\sigma}'=\text{vec}\,(\boldsymbol{I})$, i.e.,

$$\begin{aligned}\lim_{i\to\infty}E\left[e_a^2(i)\right]&=\mu^2\sigma_v^2 E\left[\|\overline{\mathbf{u}}_i\|^2_{(I-F)^{-1}\text{vec}(R)}\right]\\ \lim_{i\to\infty}E\left[\|\tilde{\boldsymbol{w}}_i\|^2\right]&=\mu^2\sigma_v^2 E\left[\|\overline{\mathbf{u}}_i\|^2_{(I-F)^{-1}\text{vec}(I)}\right]\end{aligned}$$

4. THE ERROR NONLINEARITY CASE

In this case, $\boldsymbol{H}(\mathbf{u}_i)=\boldsymbol{I}$. However, the analysis is more demanding and we shall assume that the filter is long enough for the following assumptions to be reasonable:

AG $e_a(i)$ is Gaussian.

AU $\|\mathbf{u}_i\|^2$ and $f^2(e(i))$ are uncorrelated.

For long adaptive filters, the first assumption is justified by central-limit theorem arguments while the latter is a weaker version of the independence assumption (it becomes more accurate as the filter gets longer).

Thus consider relations (5) and (6) for $\boldsymbol{\Sigma}=\boldsymbol{H}(\mathbf{u}_i)=\boldsymbol{I}$. By eliminating $e_p(i)$ from both equations, we get a recursion similar to (7) for the nonlinear error case:

$$\|\tilde{\boldsymbol{w}}_{i+1}\|^2=\|\tilde{\boldsymbol{w}}_i\|^2-2\mu f(e(i))e_a(i)+\mu^2\|\mathbf{u}_i\|^2 f^2(e(i))$$

Upon taking the expectations of both sides,

$$\begin{aligned}E\left[\|\tilde{\boldsymbol{w}}_{i+1}\|^2\right]=E\left[\|\tilde{\boldsymbol{w}}_i\|^2\right]-2\mu E\left[f(e(i))e_a(i)\right]\\ +\mu^2 E\left[\|\mathbf{u}_i\|^2 f^2(e(i))\right] \quad (17)\end{aligned}$$

we see that two expectations call for evaluation. Since $e_a(i)$ is Gaussian, we have by Price theorem,

$$\begin{aligned}E\left[f(e(i))e_a(i)\right]&=E\left[e_a^2(i)\right]E[f'(e(i))]\\ &=E\left[e_a^2(i)\right]h\left(E\left[e_a^2(i)\right]\right) \quad (18)\end{aligned}$$

for some function $h(\cdot)$. By assumption AU, we can also write

$$\begin{aligned}E\left[\|\mathbf{u}_i\|^2 f^2(e(i))\right]&=E\left[\|\mathbf{u}_i\|^2\right]E\left[f^2(e(i))\right]\\ &=\text{Tr}\,(\boldsymbol{R})\;q\left(E\left[e_a^2(i)\right]\right) \quad (19)\end{aligned}$$

for some function $q(\cdot)$. Notice that in (18) and (19), $E[f'(e(i))]$ and $E[f^2(e(i))]$ depend on $e_a(i)$ through the second moment $E\left[e_a^2(i)\right]$ only, since $e_a(i)$ is Gaussian and independent of the noise. Table 2 lists the expressions for the functions $h(\cdot)$ and $q(\cdot)$ for the error nonlinearities of Table 1.

Table 2: $h(\cdot)$ and $q(\cdot)$ for the error nonlinearities of Table 1 and for Gaussian noise ($\sigma_e^2 \triangleq E[e_a^2(i)]$)

$h(\sigma_e^2)$	$q(\sigma_e^2)$
1	$\sigma_e^2 + \sigma_v^2$
$3(\sigma_e^2 + \sigma_v^2)$	$15(\sigma_e^2 + \sigma_v^2)^3$
$\frac{(2k+2)!}{2^{k+1}(k+1)!}(\sigma_e^2 + \sigma_v^2)^k$	$\frac{(4k+2)!}{2^{2k+1}(2k+1)!}(\sigma_e^2 + \sigma_v^2)^{2k+1}$
$a + 3b\sigma_e^2\sigma_v^2 + 3b\sigma_e^2$	$a^2(\sigma_e^2 + \sigma_v^2) + 6ab(\sigma_e^2 + \sigma_v^2)^2 + 15b^2(\sigma_e^2 + \sigma_v^2)^3$
$\frac{\sigma_z}{\sqrt{\sigma_z^2+\sigma_v^2+\sigma_e^2}}$	$\sigma_z^2 \sin^{-1}\left(\frac{\sigma_e^2+\sigma_v^2}{\sigma_e^2+\sigma_v^2+\sigma_z^2}\right)$
$\sqrt{\frac{2}{\pi}}\frac{1}{\sqrt{\sigma_e^2+\sigma_v^2}}$	1

To determine the steady-state performance of the algorithms, we note that in steady-state, $E\left[\|\tilde{\boldsymbol{w}}_{i+1}\|^2\right] = E\left[\|\tilde{\boldsymbol{w}}_i\|^2\right]$ as $i \to \infty$. Let $S = \lim_{i\to\infty} E\left[e_a^2(i)\right]$. Then (17) leads to

$$\boxed{S = \frac{\mu}{2}\mathrm{Tr}\,(\boldsymbol{R})\,\frac{q(S)}{h(S)}} \quad (20)$$

This expression shows that the mean-square error, S, is a fixed point of the function $\frac{\mu}{2}\mathrm{Tr}\,(\boldsymbol{R})\,\frac{q(S)}{h(S)}$. For a given error nonlinearity, we can therefore determine S by first determining h and q and then solving for S.

To study stability, we consider recursion (17) again and note that if μ is chosen to satisfy for all i:

$$\mu^2 E\left[\|\mathbf{u}_i\|^2 f^2(e(i))\right] \leq 2\mu E\left[f(e(i))e_a(i)\right]$$

then $E\left[\|\tilde{\boldsymbol{w}}_{i+1}\|^2\right] \leq E\left[\|\tilde{\boldsymbol{w}}_i\|^2\right]$, i.e., the mean-square deviation will be a decreasing and hence convergent sequence. Now by Cauchy-Schwartz inequality, we have

$$\begin{aligned} E\left[\|\mathbf{u}_i\|^2 f^2(e(i))\right] &\leq E\left[\|\mathbf{u}_i\|^4\right]^{1/2} E\left[f^4(e(i))\right]^{1/2} \\ &= E\left[\|\mathbf{u}_i\|^4\right]^{1/2} p\left(E[e_a^2(i)]\right) \end{aligned}$$

for some function $p(\cdot)$. Hence, a more conservative condition on μ for stability is

$$\mu \leq \min_{E[e_a^2(i)]} \frac{2E[e_a^2(i)]h\left(E[e_a^2(i)]\right)}{E\left[\|\mathbf{u}_i\|^4\right]^{1/2} p\left(E[e_a^2(i)]\right)} \quad (21)$$

Minimizing (21) over $E[e_a^2(i)]$ can be demanding. Instead, we know that $E[e_a^2(i)]$ is lower-bounded by the Cramer-Rao bound γ of the underlying estimation process. To obtain an upper bound, we note that if μ is chosen to satisfy (21), then

$$E\left[\|\tilde{\boldsymbol{w}}_i\|^2\right] \leq E\left[\|\boldsymbol{w}_{i-1}\|^2\right] \leq \cdots \leq E\left[\|\boldsymbol{w}_0\|^2\right]$$

Therefore, since $e_a(i)$ is Gaussian, we have

$$E\left[e_a(i)^2\right] = \frac{1}{4}\left[E|e_a(i)|\right]^2 = \frac{1}{4}E\left[|\mathbf{u}_i\tilde{\boldsymbol{w}}_i|\right]^2$$
$$\leq \frac{1}{4}E\left[\|\mathbf{u}_i\|^2\right]^{1/2} E\left[\|\tilde{\boldsymbol{w}}_i\|^2\right]^{1/2} \leq \frac{1}{4}[\mathrm{Tr}\,(\mathbf{R})]^{1/2} E\left[\|\boldsymbol{w}_0\|^2\right]^{1/2}$$

This prompts us to define the feasibility set

$$\Omega = \left\{\alpha : \gamma \leq \alpha \leq \frac{1}{4}[\mathrm{Tr}\,(\mathbf{R})]^{1/2} E\left[\|\boldsymbol{w}_0\|^2\right]^{1/2}\right\}$$

By carrying out the minimization in (21) over the set Ω, we get the following condition for stability

$$\boxed{\mu \leq \min_{\alpha\in\Omega} \frac{2\alpha\; h(\alpha)}{E\left[\|\mathbf{u}_i\|^4\right]^{1/2} p(\alpha)}} \quad (22)$$

By reviewing the above stability argument, we see that only the Gaussian assumption AG was used. Explicit bounds on μ can be obtained by evaluating h and p and carrying out the minimization in (22).

5. CONCLUSION

In this paper, we presented a unified approach for the transient analysis of adaptive filters. Among other results, we provided conditions for stability and expressions for the steady-state error.

6. REFERENCES

[1] D. L. Duttweiler, "Adaptive filter performance with nonlinearities in the correlation multiplier," *IEEE Transactions on Acoustics, Speech and Signal Processing*, vol. 30, no. 4, pp. 578-586, Aug. 1982.

[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, pp. 275-283, Aug. 1984.

[3] N. J. Bershad. "Analysis of the normalized LMS algorithm with Gaussian inputs," *IEEE Trans. Acoust. Speech Signal Process.*, vol. 34, pp. 793–806, 1986.

[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 Transactions 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 filters," *IEEE Transactions on Signal Processing,* vol. 49, no. 2, pp. 314–324, February 2001. [See also *Proc. 4th IEEE-EURASIP Workshop on Nonlinear Signal and Image Processing*, vol. 2, pp. 699–703, Antalya, Turkey, June 1999.]