A POSTERIORI ANALYSIS OF ADAPTIVE BLIND EQUALIZERS

M. Rupp[1] *and S.C. Douglas*[2]

[1]Wireless Research Laboratory, Lucent Technologies, Holmdel, NJ 07733 USA

[2]Dept. of Electrical Engineering, Southern Methodist University, Dallas, TX 75275 USA

ABSTRACT

For adaptive filters that employ training signals, the relationship between the *a priori* and *a posteriori* error signals can be used to quickly and easily characterize the stability and robustness of any given adaptive algorithm. In this paper, we extend these analytical techniques to Bussgang-type and Godard/CMA blind equalizers. We derive conditions on the nonlinearity and step size parameter to guarantee L_2 stability of the adaptive filter coefficients, and we study the resulting behavior of the algorithms in a deterministic context. In addition, we give conditions under which an algorithm that is based on an *a posteriori* blind error criterion is L_2 stable.

1. INTRODUCTION

Adaptive filters employing training signals are widely used in a number of practical applications. One of the most popular adaptive filters is the normalized least-mean-square (NLMS) adaptive filter, in which the FIR filter coefficients are updated as

$$\mathbf{w}(k+1) = \mathbf{w}(k) + \frac{\overline{\mu}(k)}{||\mathbf{x}(k)||^2} e(k)\mathbf{x}^*(k) \quad (1)$$

$$e(k) = d(k) - y(k) \quad (2)$$

$$y(k) = \mathbf{x}^T(k)\mathbf{w}(k), \quad (3)$$

where $\mathbf{w}(k) = [w_0(k) \ \cdots \ w_{L-1}(k)]^T$ and $\mathbf{x}(k) = [x(k) \cdots x(k-L+1)]^T$ are the coefficient and input signal vectors at time k, respectively, $e(k)$ is the *a priori* error signal at time k, $||\mathbf{x}(k)||^2$ denotes the L_2-norm of the vector $\mathbf{x}(k)$, $\overline{\mu}(k)$ is a step size parameter, and $*$ denotes complex-conjugate. Various analytical studies of the NLMS adaptive filter have been provided in the literature, in which its stability and robustness properties have been elucidated [1]–[11]. Recently, a novel analytical method employing the relationship between the *a posteriori* error signal

$$e_p(k) \stackrel{\Delta}{=} d(k) - \mathbf{x}^T(k)\mathbf{w}(k+1) \quad (4)$$

and the *a priori* error signal in (2) has been developed and applied to a wide class of adaptive filtering algorithms [12]. This technique provides a simple way of characterizing the stability and robustness of adaptive filters that employ training signals without statistical assumptions on the signals being processed.

[0]This research was supported by NSF Grant No. MIP-9501680.

In some adaptive filter applications, it may be undesirable to employ a training signal $d(k)$. For example, in wireless communication from a basestation to multiple receivers, having the basestation transmit a training signal to each receiver reduces the effective data rate at which information can be broadcast and complicates the transmission protocol. For this reason, blind adaptive algorithms have been developed that can, in certain circumstances, acquire a desirable solution for $\mathbf{w}(k)$ without a training signal. A large class of such blind algorithms can be expressed in the general form

$$\mathbf{w}(k+1) = \mathbf{w}(k) + \mu(k)y(k)g(y(k))\mathbf{x}^*(k), \quad (5)$$

where $g(y)$ is an even function that depends on the nature of the received signal $x(k)$. For example, if

$$g(y) = |y|^{q-2}[A^q - |y|^q]^{p-2}[A^q - |y|^q] \quad (6)$$

for some integers $(p, q) \geq 1$ and real-valued $A > 0$, then (5) is the Godard blind equalizer [13], and if $p = q = 2$, the constant modulus algorithm (CMA) is obtained [14]. Generally, the analysis of blind algorithms is much more challenging than algorithms employing training signals. In particular, developing conditions on the step size $\mu(k)$ and nonlinearity $g(y)$ to guarantee the stability of $\mathbf{w}(k)$ is particularly problematic. While some statistical analyses have been provided in the literature [15]–[17], few deterministic methods have been developed [9]–[10].

In this paper, we provide a deterministic analysis of blind adaptive algorithms of the form in (5). Unlike other methods [15]–[17], ours yields results that hold for any bounded-magnitude received signal $x(k)$. We develop conditions on $\mu(k)$ and $g(y)$ that guarantee the L_2 stability of the filter coefficients and of the equalizer output sequence. Focusing on the Godard/CMA family, we then show that the algorithm that provides zero *a posteriori* error is normalized CMA [18]–[21]. Since NLMS and other zero-*a-posteriori*-error algorithms are known to have fast convergence properties, our results provide some justification for the use of normalized CMA. We also examine *a posteriori* Godard/CMA forms and explore their stability properties.

2. A POSTERIORI ANALYSIS FOR TRAINED ADAPTATION

We first review the *a posteriori* analysis results in [12] to provide a point of comparison with the newly-derived results for blind adaptation in the next section. In this case, all quantities are assumed to be real-valued.

1. For the trained algorithm

$$\mathbf{w}(k+1) = \mathbf{w}(k) + [e(k) - Q(e(k))]\frac{\mathbf{x}(k)}{||\mathbf{x}(k)||^2} \quad (7)$$

where $Q(e)$ is any function, the *a posteriori* error is

$$e_p(k) = Q(e(k)). \quad (8)$$

2. Let $d(k)$ be of the form

$$d(k) \quad = \quad \mathbf{x}^T(k)\mathbf{w}_{opt} + \eta(k), \quad (9)$$

where $\mathbf{w}_{opt}$ is an optimum coefficient vector, $\eta(k)$ is L_2-bounded, and $||\mathbf{x}(k)||^2 \geq \delta > 0$ for all k. Define the noiseless error $\overline{e}(k)$ as

$$\overline{e}(k) \quad \triangleq \quad e(k) - \eta(k). \quad (10)$$

Then, if $Q(e)$ is contractive such that

$$|e_p(k)| \leq \beta(k)|e(k)| \quad (11)$$

for some

$$0 \ \leq \ \beta(k) \ < \ 1, \quad (12)$$

then the algorithm is L_2-stable and satisfies

$$\lim_{n\to\infty} \frac{1}{n}\sum_{k=0}^{n-1} \frac{\overline{e}^2(k)}{||\mathbf{x}(k)||^2} = K \lim_{n\to\infty} \frac{1}{n}\sum_{k=0}^{n-1} \frac{\eta^2(k)}{||\mathbf{x}(k)||^2} \quad (13)$$

for some $0 < K < \infty$.

3. Setting $\beta(k) = 0$ in (11) yields the NLMS algorithm in (1) with $\overline{\mu}(k) = 1$, an algorithm that is known to provide fast adaptation behavior.
4. The *a posteriori* update given by

$$\mathbf{w}(k+1) \quad = \quad \mathbf{w}(k) + \mu_p(k)f(e_p(k))\mathbf{x}(k) \quad (14)$$

with $f(e)$ being any odd-symmetric monotonically-increasing function is L_2-stable in the sense of (9)–(13) for any $\mu_p(k) > 0$

5. Letting $\mu_p(k) \to \infty$ for (14) yields the NLMS algorithm in (1) with $\overline{\mu}(k) = 1$.

3. A POSTERIORI ANALYSIS FOR BLIND ADAPTATION

We now consider the behavior of the blind adaptive algorithm in (5) from a deterministic standpoint. For this analysis, we have several choices to make regarding the nature of the performance evaluation. Perhaps the most important is the choice of error criterion from which to evaluate the algorithm's performance and behavior. Here, we take the property-restoral viewpoint that motivates the use of Godard/CMA adaptation in blind equalizers, which states that $|y(k)|$ should tend to a constant factor A over time. Thus, we choose as our error function the quantity

$$e(k) = A^2 - |y(k)|^2. \quad (15)$$

If $|e(k)|$ tends to small values over time, then we shall say that the algorithm has properly converged. Note that such a condition does not necessarily guarantee that $\mathbf{w}(k)$ converges to a desirable solution, as this more-stringent condition depends on the nature of $x(k)$.

3.1. L_2 Output Stability of Adaptive Blind Equalizers

As in the trained adaptation case, define

$$e_p(k) \triangleq A^2 - |y_p(k)|^2, \quad (16)$$

where the *a posteriori* output is defined as

$$y_p(k) \quad \triangleq \quad \mathbf{x}^T(k)\mathbf{w}(k+1). \quad (17)$$

Using (5), the relationship between $y_p(k)$ and the *a priori* output $y(k)$ is

$$y_p(k) \quad = \quad [1 + \alpha(k)g(y(k))]y(k) \quad (18)$$

$$\alpha(k) \quad \triangleq \quad \mu(k)||\mathbf{x}(k)||^2. \quad (19)$$

Substituting this relationship into (16), we obtain

$$e_p(k) \quad = \quad (1 - a(k))e(k) \quad (20)$$

$$a(k) \quad \triangleq \quad \frac{\alpha(k)|y(k)|^2 g(y(k))}{A^2 - |y(k)|^2}[2 + \alpha(k)g(y(k))]. \quad (21)$$

Define

$$\beta(k) \quad \triangleq \quad |1 - a(k)|. \quad (22)$$

We now state and prove the following theorem.

Theorem 1: *If $g(y)$ and $\alpha(k)$ are chosen such that (11)–(12) are true for the definitions in (15)–(22), then (5) is L_2-stable and produces an output sequence $y(k)$ that satisfies*

$$\lim_{n\to\infty} \frac{1}{n}\sum_{k=0}^{n-1} a(k)\frac{|y(k)|^2}{||\mathbf{x}(k)||^2} = \lim_{n\to\infty} \frac{1}{n}\sum_{k=0}^{n-1} a(k)\frac{A^2}{||\mathbf{x}(k)||^2} \quad (23)$$

with $0 < a(k) < 2$.

Proof: To prove the theorem, we first show that the magnitude of $||\mathbf{w}(k)||^2$ remains bounded over time. Taking the squared L_2 norms of both sides of (5) yields

$$||\mathbf{w}(k+1)||^2 = ||\mathbf{w}(k)||^2 + \frac{\alpha(k)|y(k)|^2 g(y(k))}{||\mathbf{x}(k)||^2}[2 + \alpha(k)g(y(k))] \quad (24)$$

$$= ||\mathbf{w}(k)||^2 + \frac{e(k) - e_p(k)}{||\mathbf{x}(k)||^2}, \quad (25)$$

where we have used the relationships in (15)–(19) in going from (24) to (25). Noting (20), we have

$$||\mathbf{w}(k+1)||^2 \quad = \quad ||\mathbf{w}(k)||^2 + a(k)\frac{e(k)}{||\mathbf{x}(k)||^2} \quad (26)$$

or

$$||\mathbf{w}(k+1)||^2 + a(k)\frac{|y(k)|^2}{||\mathbf{x}(k)||^2} = ||\mathbf{w}(k)||^2 + a(k)\frac{A^2}{||\mathbf{x}(k)||^2}. \quad (27)$$

From the Cauchy-Schwartz inequality, we have

$$|y(k)|^2 \quad = \quad \gamma(k)||\mathbf{x}(k)||^2\,||\mathbf{w}(k)||^2 \quad (28)$$

where $0 \leq \gamma(k) \leq 1$, and $0 < \gamma(k) \leq 1$ if we ignore updates for which $\mathbf{x}(k)$ and $\mathbf{w}(k)$ are orthogonal. Thus,

$$||\mathbf{w}(k+1)||^2 = (1-\gamma(k)a(k))||\mathbf{w}(k)||^2 + a(k)\frac{A^2}{||\mathbf{x}(k)||^2}. \quad (29)$$

This recursion is L_2-bounded so long as $|1-\gamma(k)a(k)| < 1$. Since $0 < \gamma(k) \leq 1$, a sufficient condition on $\beta(k)$ in (22) to guarantee stability of $||\mathbf{w}(k)||^2$ stability is that given in (12).

To prove the second part of the theorem, we iterate (27) from $0 \leq k \leq n-1$ to obtain

$$\begin{aligned} &||\mathbf{w}(n)||^2 + \sum_{k=0}^{n-1} a(k)\frac{|y(k)|^2}{||\mathbf{x}(k)||^2} \\ &= ||\mathbf{w}(0)||^2 + \sum_{k=0}^{n-1} a(k)\frac{A^2}{||\mathbf{x}(k)||^2}. \end{aligned} \quad (30)$$

Since $||\mathbf{w}(n)||^2$ is bounded, we divide both sides of the above equation by n and take limits as n approaches ∞, which yields (23) □.

3.2. Implications of Analysis

We can make several remarks regarding the results of Theorem 1.

Remark #1: The condition in (23) is somewhat weaker than that in (13) for the trained case. Eqn. (23) implies that the blind adaptive algorithm can provide proper gain control of the output signal $y(k)$ and can be compared to "convergence in the mean" in a statistical environment. Even so, this result is independent of the nature of $x(k)$. Obtaining stronger stability conditions would require assuming a model structure for $x(k)$ such that there exists a $\mathbf{w}(k)$ for which $|y(k)| \approx A$. For an analysis of this type, see [9]–[10].

Remark #2: The condition on $\beta(k)$ in (12) can be expressed in terms of $a(k)$ in (21) as

$$0 < a(k) < 2. \quad (31)$$

Similar bounds exist for $\alpha(k)$ in the trained adaptation case. Moreover, (23) is similar in structure to the bound in the trained adaptation case for $\alpha(k) = a(k)$. Thus, $a(k)$ for (5) plays the role of $\alpha(k)$ in the trained adaptation case; *i.e.* values close to zero provide slower adaptation but more-accurate gain control, whereas $a(k) = 1$ (equiv. $\beta(k) = 0$) provides fast adaptation but worse data averaging within the updates. Interestingly, the choice $a(k) = 1$ yields

$$\mu(k) = \frac{1}{||\mathbf{x}(k)||^2} \quad (32)$$

$$g(y) = \frac{A}{|y|} - 1, \quad (33)$$

such that normalized CMA is obtained. Since normalized CMA is known to provide fast adaptation behavior, such a result is consistent with previous results on *a posteriori* error algorithms [6, 12, 18]–[21]. In fact, we can choose $a(k)$ to be a certain sequence satisfying (12) and obtain the form of $g(y)$ from (22) for this choice. An example below illustrates this calculation.

Remark #3: The condition in (12) implies that $\mu(k)$ and $g(y(k))$ must satisfy certain constraints for all k. Define

$$h(y) \triangleq \frac{g(y)}{A^2 - |y|^2}. \quad (34)$$

Since $a(k)$ in (21) must be positive, we have for $\mu(k) > 0$ that

$$g(A) = 0 \quad (35)$$

and

$$h(y) \geq 0 \quad (36)$$

for all y, with the unique possibility that $h(A) = 0$. Several choices of $g(y)$ are therefore possible, including

$$g(y) = \log(A) - \log(|y|) \quad (37)$$

$$g(y) = -1 + \sqrt{\frac{1}{2}\left(\frac{A^2}{|y|^2} + 1\right)} \quad (38)$$

$$g(y) = \frac{A - |y|^q}{A + |y|^q}, \quad q \geq 1, \quad (39)$$

and (6). Different choices for $g(y)$ clearly lead to different algorithm behavior depending on the underlying structure of $x(k)$. Once $g(y)$ has been selected, the upper bound on $\mu(k)$ guaranteeing (12) can then be calculated. An example below illustrates this calculation.

Alternatively, if a $\mu(k)$ is desired that does not depend on $y(k)$ or need not be bounded by a function of $y(k)$ for stability, then $g(y)$ must satisfy the boundedness conditions

$$|g(y)| \leq M \quad (40)$$

and

$$h(y) \leq N \quad (41)$$

for finite values of M and N, in addition to (35) and (36). Note that the common choice in (6) for $g(y)$ generally does not satisfy (40)–(41), whereas (39) does satisfy (40)–(41).

3.3. Examples

Example #1: Consider $g(y)$ in (6) with $p = q = 2$, such that standard CMA is obtained. Substituting this expression into (21), $a(k)$ is found to be

$$a(k) = \alpha(k)|y(k)|^2[2 + \alpha(k)\{A^2 - |y(k)|^2\}]. \quad (42)$$

The bounds in (12) can then be expressed as

$$-1 < [1 - \alpha(k)|y(k)|^2]^2 - A^2\alpha^2(k)|y(k)|^2 < 1. \quad (43)$$

Eqn. (43) is similar to that found for a statistical analysis of a blind decorrelation algorithm [22], and thus we can use the results of [[22], Appendix A] directly. We have stability of CMA if

$$0 < \mu(k) < \frac{\overline{\mu}_{max}(y(k), A)}{||\mathbf{x}(k)||^2}, \quad (44)$$

where

$$\overline{\mu}_{max}(y, A) = \begin{cases} \dfrac{1}{A^2 - |y|^2}\left(1 - \sqrt{\dfrac{2A^2}{|y|^2} - 1}\right) & \text{if } |y| \leq A\sqrt{2} \\ \dfrac{2}{|y|^2 - A^2} & \text{if } |y| > A\sqrt{2} \end{cases} \quad (45)$$

This condition can be used to test and adjust the value of $\mu(k)$ at each iteration, if necessary, to ensure the system's stability □.

Example #2: Suppose the choice $a(k) = a_0$ in (20) is desired. Substituting this choice into (21) yields the quadratic equation

$$|\alpha(k)g(y(k))|^2 + 2\alpha(k)g(y(k)) + a_0\left[1 - \frac{A^2}{|y(k)|^2}\right] = 0 \quad (46)$$

which has the positive solution

$$\alpha(k)g(y(k)) = -1 + \sqrt{a_0\left(\frac{A^2}{|y(k)|^2} - 1\right) + 1}. \quad (47)$$

If $a_0 = 1$, then normalized CMA with $\mu(k)$ and $g(y)$ given by (32) and (33) are obtained, respectively, whereas if $a_0 = 0.5$, then $\mu(k)$ and $g(y)$ are given by (32) and (38), respectively.

3.4. L_2 Error Stability of Adaptive Blind Equalizers

Although useful, Theorem 1 indicates only that a blind adaptive algorithm can achieve proper output gain control. A more-desirable result would place bounds on the sequence of normalized squared errors $|e(k)|^2/||\mathbf{x}(k)||^2$, as in the trained adaptation case. In this section, we develop such an analysis, which yields additional restrictions on the step size $\mu(k)$.

Theorem 2: *Suppose that $g(y)$ and $\alpha(k)$ are chosen such that Theorem 1 is true. If in addition $g(y)$, $h(y)$ in (34), and $\alpha(k)$ satisfy (40), (41), and*

$$0 < \alpha(k) < \frac{2}{|y(k)|^2 h(y(k))} \quad (48)$$

for all k, respectively, then (5) produces an error sequence $e(k)$ that is L_2-bounded and satisfies

$$\sum_{k=0}^{n-1} \frac{\alpha(k)h(y(k))}{||\mathbf{x}(k)||^2}[2 - \alpha(k)|y(k)|^2 h(y(k))]|e(k)|^2 \leq ||\mathbf{w}(0)||^2 + \sum_{k=0}^{n-1} \frac{2\alpha(k)}{||\mathbf{x}(k)||^2} A^2 |y(k)|^2 g(y(k)). \quad (49)$$

Proof: To prove the theorem, we consider the update equation in (5). Taking the squared L_2-norms of both sides gives

$$\begin{aligned} ||\mathbf{w}(k+1)||^2 &= ||\mathbf{w}(k)||^2 + \frac{\alpha^2(k)}{||\mathbf{x}(k)||^2}|y(k)|^2 g^2(y(k)) \\ &\quad + \frac{2\alpha(k)}{||\mathbf{x}(k)||^2}|y(k)|^2 g(y(k)) \quad (50) \\ &= ||w(k)||^2 + \frac{\alpha^2(k)}{||\mathbf{x}(k)||^2}|y(k)|^2 g^2(y(k)) \\ &\quad + \frac{2\alpha(k)}{||\mathbf{x}(k)||^2}(|y(k)|^2 - A^2) g(y(k)) \\ &\quad + \frac{2\alpha(k)}{||\mathbf{x}(k)||^2}|y(k)|^2 A^2 g(y(k)). \quad (51) \end{aligned}$$

Assume that $h(y(k))$ satisfies (36). Then, we can rewrite (51) as

$$\begin{aligned} ||\mathbf{w}(k+1)||^2 &= ||\mathbf{w}(k)||^2 \\ &\quad + \frac{\alpha^2(k)}{||\mathbf{x}(k)||^2}|y(k)|^2 h^2(y(k))[A^2 - |y(k)|^2]^2 \\ &\quad - \frac{2\alpha(k)}{||\mathbf{x}(k)||^2} h(y(k))[A^2 - |y(k)|^2]^2 \\ &\quad + \frac{2\alpha(k)}{||\mathbf{x}(k)||^2}|y(k)|^2 A^2 g(y(k)), \quad (52) \end{aligned}$$

or, equivalently,

$$\begin{aligned} ||\mathbf{w}(k+1)||^2 &+ \frac{\alpha(k)h(y(k))}{||\mathbf{x}(k)||^2}[2 - \alpha(k)|y(k)|^2 h(y(k))]|e(k)|^2 \\ &= ||\mathbf{w}(k)||^2 + \frac{2\alpha(k)}{||\mathbf{x}(k)||^2}|y(k)|^2 A^2 g(y(k)). \quad (53) \end{aligned}$$

Iterating this relation from $k = 0$ to $k = n-1$ produces (49). Now, if (48) is true, then all of the factors premultiplying $|e(k)|^2/||\mathbf{x}(k)||^2$ on the LHS of (49) are positive, and if Theorem 1 and (40) are true, then the RHS of (49) is finite. Taken together, these two facts prove the theorem □.

Remark #4: The boundedness condition in (40) is a critical component of the above theorem. Thus, algorithms in which $g(y)$ is bounded above are inherently more-robust than those for which $g(y)$ does not satisfies such a bound.

As an example, we have simulated the behavior of (5) with the choices in (6) (*e.g.* CMA) and (39), respectively, for $g(y)$ with $p = q = 2$ and under different coefficient initializations, channels, and receiver noise level conditions. Our qualitative observations are that

- the bounded nonlinearity in (39) appears to provide more-uniform convergence behavior across different channels and noise conditions for a fixed value of $\mu(k)$ than does CMA, and
- both algorithms perform well when operating in their robust step size regimes as determined by $0 < a(k) < 2$.

In other words, choosing a bounded $g(y)$ seems to yield an algorithm that requires less tuning of the step size parameter $\mu(k)$ to obtain good performance for a wide variety of conditions. These issues are the subject of on-going study.

4. A POSTERIORI OUTPUT ADAPTATION

For trained adaptation, it has been shown that algorithms that employ the *a posteriori* error directly within the coefficient updates provide guaranteed stability behavior independent of the value of $\mu(k) > 0$. Given the parallels between the trained and blind adaptive cases indicated above, it is reasonable to consider blind adaptive algorithms that employ the *a posteriori* output signal $y_p(k)$ directly within the coefficient updates. Do such algorithms possess the nice stability properties of their trained counterparts? To this end, we provide the following theorem:

Theorem 3: *If a blind adaptive algorithm can be written in the form*

$$\mathbf{w}(k+1) = \mathbf{w}(k) + \mu(k)y_p(k)g(y_p(k))\mathbf{x}^*(k), \quad (54)$$

where the a posteriori output $y_p(k)$ is as defined in (17), $g(y)$ and $h(y)$ satisfy the conditions in (35)–(36), and $\mu(k)$ is chosen such that (54) is angle-preserving, i.e.

$$\angle y_p(k) = \angle y(k), \tag{55}$$

then (54) is L_2 stable in the sense of Theorem 1.

Proof: Pre-multiplying both sides of (54) by $\mathbf{x}^T(k)$ and using the definitions in (15)–(19), we obtain

$$\left(1 - \alpha_p(k)\frac{g(y_p(k))}{A^2 - |y_p(k)|^2}e_p(k)\right) y_p(k) \quad = \quad y(k) \tag{56}$$

where we have defined $\alpha_p(k) \triangleq \mu_p(k)||\mathbf{x}(k)||^2$. If (54) is angle-preserving, then

$$\left(1 - \alpha_p(k)\frac{g(y_p(k))}{A^2 - |y_p(k)|^2}e_p(k)\right) |y_p(k)| \quad = \quad |y(k)| \tag{57}$$

Consider the case where $|y_p(k)| = A + \delta$, $\delta > 0$. Then, $e_p(k) = -2A\delta - \delta^2 < 0$, and from (36), $h(y_p(k)) > 0$. Thus, the factor premultiplying $|y_p(k)|$ on the LHS of (57) is strictly greater than one, such that $e_p(k) < e(k) < 0$. Similarly, consider the case where $|y_p(k)| = A - \delta$, $0 < \delta < A$. Then, $e_p(k) = 2A\delta - \delta^2 > 0$. Thus, the factor premultiplying $|y_p(k)|$ on the LHS of (57) is strictly less than one, such that $e(k) > e_p(k) > 0$. Combining these two results, we have a guaranteed contraction from $|e(k)|$ to $|e_p(k)|$, and thus (11)–(12) are satisfied □.

Remark #5: As in the trained adaptation case, (54) must be placed in the form of a true update for implementation purposes, such that $y_p(k)$ no longer appears. As typical choices of $g(y)$ are highly nonlinear, such a problem can be difficult to solve. For example, an *a posteriori* form of CMA involves the solution of a a cubic equation in $|y_p(k)|$ at each time instant, which is a computationally-challenging task. Hence, the *a posteriori* algorithm forms appear to be less practical in the blind adaptation case as compared to the trained case. Note that if the normalized CMA nonlinearity in (33) is chosen for $g(y)$, then the resulting algorithm is (5) with

$$\mu(k) \quad = \quad \frac{\mu_p(k)}{1 + \mu_p(k)||\mathbf{x}(k)||^2}, \tag{58}$$

a result identical to that of the trained case [6].

5. CONCLUSIONS

In this paper, we have explored the use of *a posteriori* analysis methods for blind adaptive equalizers. We have derived conditions on the algorithm update form and the step size to guarantee L_2 stability of the filter coefficients, the output sequence, and the error sequence as defined in (15). Our result provide additional justification for the use of *a posteriori*-error-based schemes such as normalized CMA by indicating in what ways their behaviors are robust. In addition, our results suggest new ways of designing algorithms to provide additional levels of robustness. The theoretical results provide a nice connection with existing results in the trained adaptation case [12] and complement statistical analyses of these schemes [15]–[17].

REFERENCES

[1] J. Nagumo and A. Noda, "A learning method for system identification," *IEEE Trans. Automat. Contr.*, vol. AC-12, pp. 282-287, June 1967.

[2] G.C. Goodwin and K.S. Sin, *Adaptive Filtering, Prediction, and Control* (Englewood Cliffs, NJ: Prentice-Hall, 1984).

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

[4] M. Tarrab and A. Feuer, "Convergence and performance analysis of the normalized LMS algorithm with uncorrelated Gaussian data," *IEEE Trans. Inform. Theory*, vol. 34, pp. 680-691, July 1988.

[5] D.T.M. Slock, "On the convergence behavior of the LMS and the normalized LMS algorithms," *IEEE Trans. Signal Processing*, vol. 41, pp. 2811-2825, Sept. 1993.

[6] S.C. Douglas and T.H.-Y. Meng, "Normalized data nonlinearities for LMS adaptation," *IEEE Trans. Signal Processing*, vol. 42, pp. 1352-1365, June 1994.

[7] M. Rupp and A.H. Sayed, "A time-domain feedback analysis of filtered-error adaptive gradient algorithms," *IEEE Trans. Signal Processing*, vol. 44, pp. 1428-1440, June 1996.

[8] A.H. Sayed and M. Rupp, "Error energy bounds for adaptive gradient algorithms," *IEEE Trans. Signal Processing*, vol. 44, pp. 1982-1989, Aug. 1996.

[9] A.H. Sayed and M. Rupp, "Robustness and convergence of adaptive schemes in blind equalization and neural tetwork training," in *Proc. Eusipco Signal Processing*, pp. 1401-1404, Sept. 1996.

[10] M. Rupp and A.H. Sayed, "Robustness and convergence of adaptive schemes in blind equalization," *Proc. 30th Asilomar Conf. Signals, Syst., Comput.*, Pacific Grove, CA, vol. 1, pp. 271-275, Nov. 1996.

[11] A.H. Sayed and M. Rupp, "An l_2-stable feedback structure for nonlinear H^∞ adaptive filtering," *Automatica*, vol. 33, no. 1, pp. 13-30, Jan. 1997.

[12] S.C. Douglas and M. Rupp, "*A posteriori* updates for adaptive filters," *Proc. 31st Asilomar Conf. Signals, Syst., Comput.*, Pacific Grove, CA, vol. 2, pp. 1641-1645, Nov. 1996.

[13] D.N. Godard, "Self-recovering equalization and carrier tracking in two-dimensional data communication systems," *IEEE Trans. Comm.*, vol. 28, pp. 1867-1875, Nov. 1980.

[14] J.R. Treichler and B.G. Agee, "A new approach to multipath correction of constant modulus signals," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 31, pp. 349-372, Apr. 1983.

[15] V. Weerackody, S.A. Kassam, and K.R. Laker, "Convergence analysis of an algorithm for blind equalization," *IEEE Trans. Commun.*, vol. 39, pp. 856-865, June 1991.

[16] R. Cusani and A. Laurenti, "Convergence analysis of the CMA blind equalizer," *IEEE Trans. Commun.*, vol. 43, part 3, pp. 1304-1307, Feb.-Mar.-Apr. 1995.

[17] Y. Li and K.J.R. Liu, "Static and dynamic convergence behavior of adaptive blind equalizers," *IEEE Trans. Signal Processing*, vol. 44, pp. 2736-2745, Nov. 1996.

[18] K. Hilal and P. Duhamel, "A convergence study of the constant modulus algorithm leading to a normalized-CMA and block-normalized CMA," *Proc. EUSIPCO*, 1992.

[19] D.L. Jones, "A normalized constant-modulus algorithm," *Proc. 29th Asilomar Conf. Signals, Syst., Comput.*, Pacific Grove, CA, vol. 1, pp. 694-697, Nov. 1996.

[20] C.B. Papadias and D.T.M. Slock, "Normalized sliding-window constant modulus and decision-directed algorithms: a link between blind equalization and classical adaptive filtering," *IEEE Trans. Signal Processing*, vol. 45, pp. 231-235, Jan. 1997.

[21] A.L. Swindlehurst, "Normalized adaptvie decision directed equalization," *IEEE Signal Processing Lett.*, vol. 5, pp. 18-19, Jan. 1998.

[22] S.C. Douglas and A. Cichocki, "Neural networks for blind decorrelation of signals," *IEEE Trans. Signal Processing*, vol. 45, pp. 2829-2842, Nov. 1997.