

Figure 3: Evolution of MSE: Feintuch's algorithm (dashed line), Algorithm #1 (dotted line), and Algorithm #2 (solid line), $\alpha_f = 0.05$, $\alpha_1 = \alpha_2 = 0.5$.

where $\hat{d}(i) = \mathbf{w}_i^T \hat{\mathbf{u}}_i$ and $\hat{\mathbf{u}}_i = [\hat{d}(i-1) \cdots \hat{d}(i-N) x(i) \cdots x(i-M+1)]^T$, respectively.

Figure 3 depicts the evolution of the MSEs for the three adaptive filters on these signals, where $\mathbf{w}_0 = [0 \ 0 \ 0]^T$ for each algorithm. Because the unknown system is not SPR, Feintuch's algorithm fails to estimate the output of the system, whereas both of the proposed algorithms provide excellent convergence characteristics in this case. These results indicate that the positiveness of $c^{(1)}(i)$ and the strictly-positive-real nature of $C_i(e^{-j\omega})$ are maintained for Algorithms #1 and #2, respectively, for all time instants for the choices of initial coefficients and step size values. These results also indicate that the proposed algorithms are not restricted for use with only a subclass of possible IIR systems, such as SPR systems.

We now repeat the above experiment for an initial coefficient vector of $\mathbf{w}_0 = [-1 \ -0.25 \ 1]^T$. Figure 2 shows the behaviors of the instantaneous squared errors for Algorithms #1 and #2, in which it is seen that the behaviors of the proposed algorithms are quite different in this case. Although Algorithm #1 provides convergence of the coefficients to their optimum values, the initial condition chosen for these experiments yields $c^{(1)}(0) < 0$, and thus this adaptive filter exhibits divergent behavior initially. However, the coefficients eventually fall into a state whereby $c^{(1)}(i) > 0$, and the algorithm becomes convergent again. Our choice of initial coefficients causes $C_0(e^{-j\omega})$ for Algorithm #2 to be non-SPR, however, and this second adaptive filter never recovers from its divergent condition. These results are observable even for smaller step sizes, indicating that the behaviors are not due to a large step size choice.

4. CONCLUSIONS

We have analyzed the robustness of two new algorithms for unbiased adaptive FIR filtering in the presence of zero-mean output noises. Our results indicate that the algorithms are locally asymptotically stable about the optimum coefficient solutions. However, the stabilities of the algorithms depend on both the unknown system and the state of the adaptive systems, and it is possible to obtain divergent behavior from each algorithm for certain combinations of unknown sys-

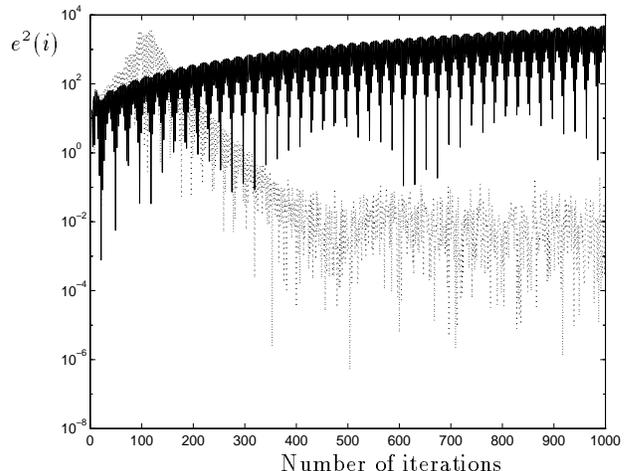


Figure 4: Evolution of $e^2(i)$ for non-zero initial coefficient values: Algorithm #1 (dotted line) and Algorithm #2 (solid line), $\alpha_1 = \alpha_2 = 0.05$.

tems and coefficient initializations. Our analysis method can be used to quantify the nature of these instabilities, allowing their effects to be carefully studied, and the results suggest that zero initial coefficient values often yield good adaptation characteristics. Simulations for a three-parameter model indicate the accuracy of the analytical results.

REFERENCES

- [1] S.C. Douglas and M. Rupp, "On bias removal and unit-norm constraints in equation-error adaptive IIR filters," presented at 30th Ann. Asilomar Conf. Sig., Syst., Comput., Pacific Grove, CA, Nov. 1996.
- [2] P.A. Regalia, "An unbiased equation error identifier and reduced-order approximations," *IEEE Trans. Signal Processing*, vol. 42, no. 6, pp. 1297-1412, June 1994.
- [3] M. Rupp and A.H. Sayed, "A time-domain feedback analysis of filtered-error adaptive gradient algorithms," *IEEE Trans. Signal Processing*, vol. 44, no. 6, pp. 1428-1440, June 1996.
- [4] A.H. Sayed and M. Rupp, "Error energy bounds for adaptive gradient algorithms," *IEEE Trans. Signal Processing*, vol. 44, no. 8, pp. 1982-1989, Aug. 1996.
- [5] B. Widrow *et al.*, "Stationary and nonstationary learning characteristics of the LMS adaptive filter," *Proc. IEEE*, vol. 64, no. 8, pp. 1151-1162, Aug. 1976.
- [6] M. Rupp, "The behavior of LMS and NLMS algorithms in the presence of spherically invariant processes," *IEEE Trans. Signal Processing*, vol. 41, no. 3, pp. 1149-1160, Mar. 1993.
- [7] S.C. Douglas and W. Pan, "Exact expectation analysis of the LMS adaptive filter," *IEEE Trans. Signal Processing*, vol. 43, no. 12, pp. 2863-2871, Dec. 1995.
- [8] P.E. Mantay, "Convergent automatic synthesis procedures for sampled data networks with feedback," Tech. Rep. SEL-6773-1, Stanford Univ., Stanford, CA, Oct. 1964.
- [9] K. Steiglitz and L.E. McBride, "A technique for the identification of linear systems," *IEEE Trans. Autom. Control*, vol. 10, pp. 461-464, 1965.
- [10] P.L. Feintuch, "An adaptive recursive LMS filter," *Proc. IEEE*, vol. 64, no. 11, pp. 1622-1624, Nov. 1976.
- [11] C.R. Johnson, "Adaptive IIR filtering: Current results and open issues," *IEEE Trans. Inform. Theory*, vol. 30, no. 2, pp. 237-250, Mar. 1984.
- [12] J.J. Shynk, "Adaptive IIR filtering," *IEEE Signal Processing Mag.*, vol. 6, no. 2, pp. 4-21, April 1989.
- [13] G. Long, F. Ling, and J.G. Proakis, "The LMS algorithm with delayed coefficient adaptation," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 37, no. 9, pp. 1397-1405, Sept. 1989; vol. 40, no. 1, pp. 230-232, Jan. 1992 (corrections).
- [14] E. Bjarnason, "Analysis of the filtered-X LMS algorithm," *IEEE Trans. Speech Audio Processing*, vol. 3, no. 6, pp. 504-514, Nov. 1995.

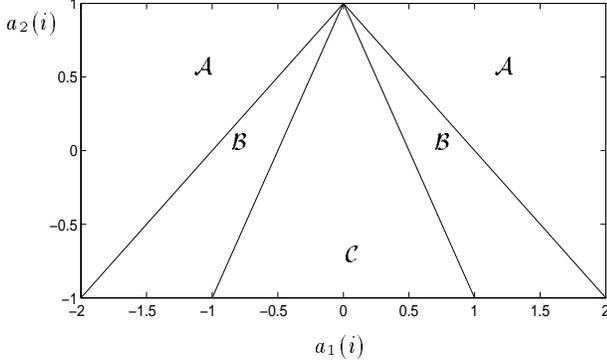


Figure 2: Contour lines for $f(a_1(i), a_2(i))$ as a function of $a_1(i)$ and $a_2(i)$: **A**: unstable parameter range, **B**: $f(a_1(i), a_2(i)) < 0$, and **C**: $f(a_1(i), a_2(i)) > 0$.

determined by

$$f(a_1(i), a_2(i)) = \min_{a_1, a_2} \frac{1 + a_1 a_1(i) + a_2 a_2(i)}{1 + a_1^2(i) + a_2^2(i)}. \quad (26)$$

For adaptive system parameter values that fall in Region **B** of this figure, there exist stable unknown systems that cause divergence of the algorithm from these parameter values. Thus, the stability of Algorithm #1 depends on both $A_i(q^{-1})$ and $A(q^{-1})$, unlike many algorithms based on output error minimization whose stabilities only depend on $A(q^{-1})$ [8]–[12].

2.2. Algorithm #2

We now consider the algorithm in (10) for $j = 2$. Applying a similar method as was used for Algorithm #1, we obtain the following relationship:

$$e(i) = v_a(i) + e_a^{(2)}(i) - c^{(2)}(i) d(i), \quad (27)$$

where $v_a(i)$ and $e_a^{(2)}(i)$ are as defined in (12)–(13) and

$$c^{(2)}(i) = \mathbf{a}_i^T (\mathbf{a}_i - \mathbf{a}). \quad (28)$$

By manipulating the input-output equations for this filter, we can also obtain the relationship

$$\begin{aligned} d(i) &= \frac{B(q^{-1})[e(i)] - B_i(q^{-1})[v_a(i)]}{B(q^{-1})(1 - A_i(q^{-1})) - B_i(q^{-1})(1 - A(q^{-1}))} \\ &= D_i(q^{-1})(B(q^{-1})[e(i)] - B_i(q^{-1})[v_a(i)]), \end{aligned} \quad (29)$$

where

$$B(q^{-1}) = \sum_{m=0}^M b_m q^{-m}, \quad B_i(q^{-1}) = \sum_{m=0}^M b_m(i) q^{-m} \quad (30)$$

$$\begin{aligned} D_i(q^{-1}) &= \frac{1}{B(q^{-1})(1 - A_i(q^{-1})) - B_i(q^{-1})(1 - A(q^{-1}))} \end{aligned} \quad (31)$$

respectively. Substituting (29) into (27) produces the relationship

$$e(i) = \frac{(1 + c^{(2)}(i) D_i(q^{-1}) B_i(q^{-1})) [v_a(i)] + [e_a^{(2)}(i)]}{1 + c^{(2)}(i) D_i(q^{-1}) B_i(q^{-1})}. \quad (32)$$

Using the results of [4], it can be shown from the above relationships that Algorithm #2 employs a filtered-error gradient update with error path filter $C_i(q^{-1})$ given by

$$C_i(q^{-1}) = \frac{1}{1 + c^{(2)}(i) D_i(q^{-1}) B_i(q^{-1})} \quad (33)$$

Thus, necessary conditions for convergence of this algorithm are (i) $C_i(q^{-1})$ is stable and (ii) $C_i(e^{-j\omega})$ is SPR at all frequencies ω that are represented in the input signal $x(i)$.

In the case where $\mathbf{a}_i \approx \mathbf{a}$ and $\mathbf{b}_i \approx \mathbf{b}$, we find from the definition of $c^{(2)}(i)$ in (28) that its value vanishes in this case. Therefore, the stability behavior of Algorithm #2 about the optimum coefficient solution is the same as that of the equation error LMS adaptive IIR filter, and the algorithm is locally asymptotically stable. In addition, by choosing $\mathbf{a}_0 = \mathbf{0}$ and $\mathbf{b}_0 = \mathbf{0}$, Algorithm #2 also behaves initially like the equation-error-based algorithm. However, during the transient adaptation phase, the stability of Algorithm #2 is no longer guaranteed. In fact, it is possible to choose \mathbf{a}_i and \mathbf{b}_i such that $C_i(e^{-j\omega})$ is not SPR for particular frequencies ω , causing divergence of the algorithm if $x(i)$ contains energy at one or more of these frequencies. Although other adaptive IIR filtering algorithms based on output error minimization also require an SPR condition, Algorithm #2 differs from these other algorithms in that the SPR condition depends on both the unknown system and the state of the adaptive system.

In addition, the stability of Algorithm #2 also depends on the step size $\mu(i)$ chosen for the algorithm. This algorithm is a filtered-error algorithm, however, and the memory of the error filter varies according to the group delay of the filter $C_i(q^{-1})$. Thus, finding sufficient conditions on $\mu(i)$ to guarantee stability of the system even when $C_i(e^{-j\omega})$ is SPR is a challenging task. The issues that govern the choice of $\mu(i)$ in this case are similar to those for the delayed LMS and filtered-X LMS adaptive algorithms [13, 14].

3. SIMULATIONS

We now explore the accuracy of our analytical results via simulation. In the following, we focus on the robustness of the adaptive algorithms in (1)–(2) and (3)–(4); for a comparison of their adaptation performance with those of other adaptive IIR filters, the reader is referred to [1].

In these examples, we generate the desired signal using an underlying IIR model with $\mathbf{a} = [1.317 \ -0.81]^T$ and $\mathbf{b} = [1]$. It can be shown that the real part of $(1 - A(e^{-j\omega}))$ for these parameter choices becomes negative for frequencies near $\omega = 1.0$. We excite this unknown system with the signal $x(n) = \sin(n)$ and add uncorrelated zero-mean Gaussian noise with variance $\sigma_v^2 = 0.0001$ to the output of this system to produce the desired response signal. Since our results suggest a normalized step size, we choose

$$\mu(i) = \frac{\alpha_j}{\|\mathbf{u}_i^{(j)}\|_2^2}, \quad (34)$$

for Algorithm # j , $j = \{1, 2\}$, where α_j for each algorithm was chosen to be small enough to provide stable behavior whenever possible. In each case, we plot either the instantaneous squared error $e^2(i)$ from one simulation run or the MSE $E\{e^2(i)\}$ as found from an average of one hundred simulation runs. For comparison, we also plot the evolution of the corresponding quantities for Feintuch's algorithm,

$$\mathbf{w}_{i+1} = \mathbf{w}_i + \frac{\alpha_f}{\|\hat{\mathbf{u}}_i\|_2^2} (d(i) - \hat{d}(i)) \hat{\mathbf{u}}_i, \quad (35)$$

for $j = \{1, 2\}$, respectively. In each case, we can express the equation error, defined as

$$e(i) = d(i) - \mathbf{w}_i^T \bar{\mathbf{u}}_i \quad (11)$$

in terms of the filtered observation noise $v_a(i)$ and the uncorrupted *a priori* error $e_a^{(j)}(i)$, defined as

$$v_a(i) = v(i) - \mathbf{a}^T \mathbf{v}_{i-1} = (1 - A(q^{-1}))[v(i)] \quad (12)$$

$$e_a^{(j)}(i) = -\tilde{\mathbf{w}}_i^T \mathbf{u}_i^{(j)}, \quad (13)$$

respectively, where

$$A(q^{-1}) = \sum_{n=1}^N a_n q^{-n} \quad (14)$$

and q^{-1} is the delay operator.

2.1. Algorithm #1

Expressing the equation error as

$$e(i) = \mathbf{w}^T \mathbf{u}_i - \mathbf{w}^T \mathbf{u}_i^{(1)} + \mathbf{w}^T \mathbf{u}_i^{(1)} - \mathbf{w}_i^T \mathbf{u}_i^{(1)} + \mathbf{w}_i^T \mathbf{u}_i^{(1)} - \mathbf{w}_i^T \bar{\mathbf{u}}_i + v(i), \quad (15)$$

we first simplify this expression using the definitions in (12)-(13) to get

$$e(i) = \frac{1}{c^{(1)}(i)} (v_a(i) + e_a^{(1)}(i)), \quad (16)$$

where

$$c^{(1)}(i) = \frac{1 + \mathbf{a}^T \mathbf{a}_i}{1 + \mathbf{a}_i^T \mathbf{a}_i}. \quad (17)$$

Secondly, by subtracting \mathbf{w} from both sides of (10), we can express the coefficient updates for Algorithm #1 in terms of the parameter error vector $\tilde{\mathbf{w}}_i$ as

$$\tilde{\mathbf{w}}_{i+1} = \tilde{\mathbf{w}}_i + \mu(i) \mathbf{u}_i^{(1)} e(i) \quad (18)$$

$$= \tilde{\mathbf{w}}_i + \bar{\mu}^{(1)}(i) \mathbf{u}_i^{(1)} (e_a^{(1)}(i) - e_a^{(1)}(i) + \frac{\mu(i)}{\bar{\mu}^{(1)}(i)} e(i)) \quad (19)$$

$$= \tilde{\mathbf{w}}_i + \bar{\mu}^{(1)}(i) \mathbf{u}_i^{(1)} (e_a(i) - \bar{v}^{(1)}(i)), \quad (20)$$

where $\bar{\mu}^{(1)}(i)$ is arbitrary and

$$\bar{v}^{(1)}(i) = \frac{\mu(i)}{c^{(1)}(i) \bar{\mu}^{(1)}(i)} v_a(i) - \left(1 - \frac{\mu(i)}{c^{(1)}(i) \bar{\mu}^{(1)}(i)}\right) e_a^{(1)}(i), \quad (21)$$

respectively. For the particular choice

$$\bar{\mu}^{(1)}(i) = \frac{1}{\|\mathbf{u}_i^{(1)}\|_2^2}, \quad (22)$$

Eqn. (20) can be shown to have the following property [3, 4]:

$$\frac{\|\tilde{\mathbf{w}}_{i+1}\|_2^2 + \bar{\mu}^{(1)}(i) e_a^{(1)2}(i)}{\|\tilde{\mathbf{w}}_i\|_2^2 + \bar{\mu}^{(1)}(i) \bar{v}^{(1)2}(i)} = 1. \quad (23)$$

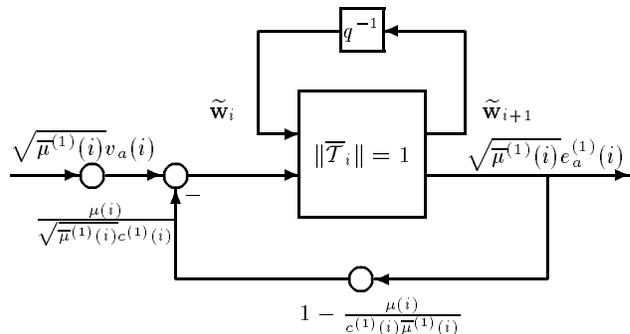


Figure 1: A time-variant lossless mapping with gain feedback for Alg. #1, uncorrelated noise case.

Equation (23) defines a lossless feedforward path mapping, whereas (21) defines a feedback path mapping. The entire structure is depicted in Figure 1, where \bar{T}_i denotes the feedforward path mapping.

Using Figure 1 and the small gain theorem, the stability of Algorithm #1 is guaranteed if the feedback gain of the system is less than one in magnitude. This condition leads to stability bounds for $\mu(i)$ as given by

$$0 < \mu(i) < \frac{2c^{(1)}(i)}{\|\mathbf{u}_i^{(1)}\|_2^2}. \quad (24)$$

Since the value of $c^{(1)}(i)$ is critical for this condition to hold, we now consider its form more carefully.

Note that near the optimum solution where $\mathbf{a}_i \approx \mathbf{a}$, we find from (17) that $c^{(1)}(i) \approx 1$, and thus the stability behavior of Algorithm #1 locally about the optimum coefficient solution is the same as that of the standard equation error LMS adaptive filter. Since the equation error LMS adaptive filter is asymptotically stable for suitably small step size values, we can conclude that Algorithm #1 is locally asymptotically stable. Moreover, without any *a priori* knowledge of the optimum coefficient values, both \mathbf{a}_0 and \mathbf{b}_0 are typically chosen to be zero vectors. Such a choice yields $c^{(1)}(0) = 1$, and thus the behavior of Algorithm #1 is initially the same as that of the standard equation error LMS adaptive filter.

The stability of Algorithm #1 is not guaranteed during its transient adaptation phase for all unknown models. From (24), we see that $c^{(1)}(i) > 0$ is necessary for a stable choice of $\mu(i)$ to exist. This condition is equivalent to having the vector $[1 \ -\mathbf{a}^T]$ lie in the half-plane defined by the vector $[1 \ -\mathbf{a}^T]$. In the frequency domain, this condition is equivalent to

$$\int_{-\pi}^{\pi} (1 - A_i(e^{-j\omega}))(1 - A_i^*(e^{-j\omega})) d\omega > 0, \quad (25)$$

where $A_i(q^{-1})$ for the adaptive system AR parameters is defined similarly to (14). It can be shown that for $N = 1$, all choices of $a_1(i)$ that yield a bounded-input, bounded-output stable adaptive filter model also satisfy (25). For $N \geq 2$, however, the stability condition can be violated for certain choices of $\{a_n(i)\}$, $1 \leq n \leq N$. Figure 2 depicts contour lines for minimal values of $c^{(1)}(i)$ for different $a_1(i)$ and $a_2(i)$ within the range of stable models for $N = 2$, as

DETERMINISTIC STABILITY ANALYSES OF UNIT-NORM CONSTRAINED ALGORITHMS FOR UNBIASED ADAPTIVE IIR FILTERING

M. Rupp[†] and *S.C. Douglas*[‡]

[†]Wireless Research Laboratory, Lucent Technologies, Crawford Hill, Holmdel, NJ 07733 USA

[‡]Department of Electrical Engineering, University of Utah, Salt Lake City, UT 84112 USA

ABSTRACT

Recently, two simple gradient-based algorithms for unbiased IIR system identification in the presence of zero-mean correlated output noise were derived and shown to perform well in simulation [1]. In this paper, we study the stability and robustness of these two adaptive filters, deriving strictly positive real (SPR) conditions on the overall unknown-plus-adaptive systems to guarantee convergence of the coefficients to their optimum values. Unlike other algorithms for unbiased IIR adaptive filtering, the stability of each of these algorithms depends on the initial values of the filter coefficients. However, near the optimum coefficient solutions, both algorithms are locally-stable, irrespective of the unknown system. Simulations verify the results of our analyses.

1. INTRODUCTION

This paper presents robustness and stability analyses of two algorithms for adaptive IIR filters [1]. These algorithms minimize the equation error cost function according to a constraint on the autoregressive parameters [2] and can provide unbiased estimates of an unknown system's coefficients for potentially-correlated output noises. The coefficient updates for these algorithms are:

Algorithm #1:

$$\mathbf{a}_{i+1} = \mathbf{a}_i + \mu(i)e(i) \left(\mathbf{d}_{i-1} + \frac{e(i) (\mathbf{R}_{vv}\mathbf{a}_i - \mathbf{p}_{vv})}{1 - 2\mathbf{p}_{vv}^T\mathbf{a}_i + \mathbf{a}_i^T\mathbf{R}_{vv}\mathbf{a}_i} \right) \quad (1)$$

$$\mathbf{b}_{i+1} = \mathbf{b}_i + \mu(i)e(i)\mathbf{x}_i, \quad (2)$$

Algorithm #2:

$$\mathbf{a}_{i+1} = \mathbf{a}_i + \mu(i)e(i) \left(\mathbf{d}_{i-1} + d(i) \frac{\mathbf{R}_{vv}\mathbf{a}_i - \mathbf{p}_{vv}}{1 - \mathbf{p}_{vv}^T\mathbf{a}_i} \right) \quad (3)$$

$$\mathbf{b}_{i+1} = \mathbf{b}_i + \mu(i)e(i)\mathbf{x}_i, \quad (4)$$

where $\mathbf{a}_i = [a_1(i) \cdots a_N(i)]^T$ and $\mathbf{b}_i = [b_0(i) \cdots b_{M-1}(i)]^T$ are the autoregressive and moving average coefficient vectors, $\mathbf{x}_i = [x(i) \cdots x(i-M+1)]^T$ and $\mathbf{d}_i = [d(i) \cdots d(i-N+1)]$ are the input and desired response signal vectors, $e(i) = d(i) - \mathbf{a}_i^T\mathbf{d}_{i-1} - \mathbf{b}_i^T\mathbf{x}_i$ is the equation error, \mathbf{R}_{vv} and \mathbf{p}_{vv} are the N -dimensional normalized autocorrelation matrix and vector of the observation noise, and $\mu(i)$ is the step size. Both of these adaptive filters assume that the desired response signal is generated from an IIR filter with parameter vectors \mathbf{a} and \mathbf{b} whose output $y(i)$ is corrupted by an additive zero mean observation noise signal $v(i)$ such that

$$d(i) = y(i) + v(i) \quad (5)$$

$$y(i) = \mathbf{a}^T\mathbf{y}_{i-1} + \mathbf{b}^T\mathbf{x}_i. \quad (6)$$

If \mathbf{R}_{vv} and \mathbf{p}_{vv} are unknown, both can be accurately estimated from signals available to the system [1]. Although statistical analyses and simulations in [1] indicate that these algorithms achieve unbiased parameter estimates for low-order system identification tasks, no formal analysis of the stability behavior of these algorithms has been given. In particular, conditions on \mathbf{a} , \mathbf{b} , \mathbf{a}_i , and \mathbf{b}_i to guarantee convergence of the algorithms have not been presented.

In this paper, we provide robustness analyses of the two adaptive algorithms in (1)-(2) and (3)-(4). Our analyses are based on a deterministic framework that has been used in [3] and [4] to determine strictly positive real (SPR) and stability conditions on the unknown system and step size, respectively, to guarantee convergence of a wide class of gradient-based adaptive filters. Unlike other analyses that are statistically-based [5]-[7], our results are independent of the statistics of the input signals. From our analysis, we show that both Algorithms #1 and #2 are both initially stable for $\mathbf{a}_0 = \mathbf{0}$ and $\mathbf{b}_0 = \mathbf{0}$ and are locally-stable about the optimum coefficient solution for suitably small step sizes and for systems that are not under-parameterized. However, during the systems' transient phases, the adaptive filters can become unstable, and we quantify the nature of this instability. Simulations verify the results of our analyses and indicate the behaviors of the adaptive filters in different situations.

2. DETERMINISTIC STABILITY ANALYSES

For brevity, we consider situations in which $v(i)$ in (5) is uncorrelated such that $\mathbf{R}_{vv} = \mathbf{I}$ and $\mathbf{p}_{vv} = \mathbf{0}$. Our analysis, however, can be extended to arbitrary correlated noises.

For these analyses, we define the vectors \mathbf{w} , \mathbf{w}_i , $\tilde{\mathbf{w}}_i$, \mathbf{u}_i , $\bar{\mathbf{u}}_i$, and $\mathbf{u}_i^{(j)}$, as

$$\mathbf{w} = \begin{bmatrix} \mathbf{a} \\ \mathbf{b} \end{bmatrix}, \quad \mathbf{w}_i = \begin{bmatrix} \mathbf{a}_i \\ \mathbf{b}_i \end{bmatrix}, \quad \tilde{\mathbf{w}}_i = \mathbf{w}_i - \mathbf{w}, \quad (7)$$

$$\mathbf{u}_i = \begin{bmatrix} \mathbf{y}_{i-1} \\ \mathbf{x}_i \end{bmatrix}, \quad \bar{\mathbf{u}}_i = \mathbf{u}_i + \begin{bmatrix} \mathbf{v}_{i-1} \\ \mathbf{0} \end{bmatrix}, \quad \mathbf{u}_i^{(j)} = \bar{\mathbf{u}}_i - \mathbf{z}_i^{(j)}, \quad (8)$$

respectively, where $\mathbf{v}_i = [v(i) \cdots v(i-N+1)]^T$ and $\mathbf{z}_i^{(j)}$ for $j = \{1, 2\}$ is given by

$$\mathbf{z}_i^{(1)} = - \left[\frac{e(i)}{1 + \mathbf{a}_i^T\mathbf{a}_i} \mathbf{a}_i \right], \quad \mathbf{z}_i^{(2)} = - \begin{bmatrix} d(i)\mathbf{a}_i \\ \mathbf{0} \end{bmatrix}, \quad (9)$$

respectively. Using these definitions, the two algorithms in (1)-(2) and (3)-(4) can be expressed as

$$\mathbf{w}_{i+1} = \mathbf{w}_i + \mu(i)e(i)\mathbf{u}_i^{(j)} \quad (10)$$