SUFFICIENT EXCITATION AND STABLE REDUCED-ORDER ADAPTIVE IIR FILTERING

C. Richard Johnson, Jr.*

Department of Electrical Engineering
Virginia Polytechnic Institute and State University
Blacksburg, VA 24061

B. D. O. Anderson**

Department of Electrical Engineering
University of Newcastle
New South Wales, 2308 Australia

ABSTRACT

A perturbed error system is developed to describe the parameter and output error behaviors of reduced-order application of adaptive identifier/filters such as the hyperstable adaptive recursive filter (HARF). Given satisfaction of a sufficient excitation condition, this error system is shown to exhibit a bounded-input, bounded-output (BIBO) property. This implies that despite order insufficiency the output and parameter estimates of HARF (and similar adaptive identifier/filter algorithms) remain bounded.

I. INTRODUCTION

In [1] the equivalence of the adaptive infinite impulse response (IIR) filtering problem and the adaptive output error identification problem solved in [2] and [3] is exploited. For asymptotic stability the algorithms of [1]-[3] require that the order of the autoregressive-moving-average (ARMA) described adaptive identifier/filter (AIF) match or exceed that of the minimal ARMA process generating the desired output. This paper addresses the reduced-order AIF question. That is, if the order of the AIF is lower than that necessary to yield no output error but for some bounded parameter values could yield a bounded output error, will the output and parameters of the AIF remain bounded?

Resolution of this robustness question is begun in the next section by relating the reduced-order application of [1] (and therefore [2]-[3]) to a perturbed error model based on the free version in [4]. The third section supports a limited affirmative answer to the boundedness question by Lyapunov function analysis. This is followed by comments on the sufficient excitation requirement of this approach to reduced-order adaptive IIR filtering analysis.

II. REDUCED-ORDER ADAPTIVE IDENTIFIER/FILTER ERROR SYSTEM

As shown in [5], the adaptive IIR filter HARF [1]

$$\hat{y}(k) = \sum_{i=1}^{n} \hat{a}_i(k)z(k-i) + \sum_{j=1}^{m} \hat{b}_j(k)u(k-j) \tag{1}$$

$$\hat{a}_i(k+1) = \hat{a}_i(k) + \mu_i z(k-i)v(k), \quad \mu_i > 0 \tag{2}$$

$$\hat{b}_j(k+1) = \hat{b}_j(k) + \rho_j u(k-j)v(k), \quad \rho_j > 0 \tag{3}$$

$$v(k) = y(k) - z(k) + \sum_{i=1}^{n} c_i[y(k-i) - z(k-i)]$$
$$= \frac{y(k) - \hat{y}(k) + \sum_{i=1}^{n} c_i[y(k-i) - z(k-i)]}{1 + \sum_{i=1}^{n} \mu_i z^2(k-i) + \sum_{j=1}^{m} \rho_j u^2(k-j)} \tag{4}$$

$$z(k) = \sum_{i=1}^{n} \hat{a}_i(k+1)z(k-i) + \sum_{j=1}^{m} \hat{b}_j(k+1)u(k-j) \tag{5}$$

with the desired output of

$$y(k) = \sum_{i=1}^{n} a_i y(k-i) + \sum_{j=1}^{m} b_j u(k-j) \tag{6}$$

can be written in the form of the error system of [4]

$$\underline{e}(k+1) = \underline{A}\underline{e}(k) + bw(k) \tag{7}$$

$$v(k) = \underline{h}^T\underline{e}(k) + dw(k) \tag{8}$$

$$w(k) = \underline{\phi}^T(k)\underline{x}(k) - \alpha\underline{x}^T(k)\underline{\Gamma}\underline{x}(k)v(k), \quad \alpha > \tfrac{1}{2}, \ \underline{\Gamma} = \underline{\Gamma}^T > 0 \tag{9}$$

$$\underline{\phi}(k+1) = \underline{\phi}(k) - \underline{\Gamma}v(k)\underline{x}(k) \tag{10}$$

given (4) and the following definitions

$$\underline{e}^T(k) \triangleq [y(k-1) - z(k-1) \quad y(k-2) - z(k-2) \ \dots \ y(k-n) - z(k-n)] \tag{11}$$

$$\underline{A} \triangleq \begin{bmatrix} a_1 & a_2 & \cdots & a_n \\ 1 & 0 & \cdots & 0 \\ 0 & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0\ 1 & 0 \end{bmatrix} \tag{12}$$

$$\underline{b}^T \triangleq [1 \quad 0 \ \dots \ 0] \tag{13}$$

$$\underline{h}^T \triangleq [c_1 + a_1 \quad c_2 + a_2 \ \cdots \ c_n + a_n] \tag{14}$$

*Supported by NSF Grant ECS-79-08695

**Supported by Australian Research Grants Comm.

$$d \triangleq 1 \tag{15}$$

$$w(k) \triangleq \sum_{i=1}^{n} [a_i - \hat{a}_i(k+1)]z(k-i) + \sum_{j=1}^{m} [b_j - \hat{b}_j(k+1)]u(k-j) \tag{16}$$

$$\underline{\phi}^T(k) \triangleq [a_1 - \hat{a}_1(k) \quad \dots \quad a_n - \hat{a}_n(k) \quad b_1 - \hat{b}_1(k) \dots b_m - \hat{b}_m(k)] \tag{17}$$

$$\underline{x}^T(k) \triangleq [z(k-1) \quad \dots \quad z(k-n) \quad u(k-1) \quad \dots u(k-m)] \tag{18}$$

$$\alpha \triangleq 1 \tag{19}$$

$$\underline{\Gamma} \triangleq \text{diag}[\mu_1 \quad \dots \quad \mu_n \quad \rho_1 \quad \dots \quad \rho_m]. \tag{20}$$

Similarly the adaptive output error identifier combining (1)-(6) in [2] and (1)-(4) in [3] can be written in the form of (7)-(10).

In the event of reduced order application, the desired output of (6) could be replaced by

$$y(k) = y_M(k) + y_U(k) \tag{21}$$

where y_M is the modeled portion of y and has the same order as the identifier filter

$$y_M(k) = \sum_{i=1}^{n} a_i y_M(k-i) + \sum_{j=1}^{m} b_j u(k-j) \tag{22}$$

and $y_U(k)$ is the unmodeled portion of y. Note that $y_U(k)$ could also represent a noise signal. A concern is whether or not z(k) and subsequently $\hat{y}(k)$ from [1] as in (1)-(5) will remain bounded if $y_U(k)$ remains bounded. Sought, using (21) and not (6) in (1)-(5), will be the implication that

$$\|y_U(k)\|^2 \leq \delta^2 < \infty \;\forall k \Rightarrow \|y(k) - z(k)\|^2 \leq \varepsilon^2 < \infty \;\forall k > \bar{k} \tag{23}$$

In [1]-[3] it was proven that $\delta^2 = 0 \Rightarrow \varepsilon^2 \to 0$. With $\delta^2 \neq 0$ it is clear that, in general, $\varepsilon^2 \neq 0$.

Given (21) and (22), (4) becomes

$$v(k) = \sum_{i=1}^{n} [c_i + a_i][y_M(k-i) - z(k-i)] + w(k) + p(k) \tag{24}$$

where w is defined in (16) and the perturbation signal p is defined by

$$p(k) \triangleq y_U(k) + \sum_{i=1}^{n} c_i y_U(k-i). \tag{25}$$

With the definition of $\underline{e}$ in (11) altered to

$$\underline{e}^T(k) \triangleq [y_M(k-1) - z(k-1) \dots y_M(k-n) - z(k-n)] \tag{26}$$

(8) would become

$$v(k) = \underline{h}^T\underline{e}(k) + dw(k) + p(k) \tag{27}$$

in order to match (24). The remainder of the error system, i.e. (7), (9), (10), and (12)-(20) would be unaltered in accommodating the effects of reduced-order modeling.

III. PERTURBED ERROR SYSTEM BIBO STABILITY

For the error model resulting from reduced-order application of [1] or [2]-[3]

$$\underline{e}(k+1) = \underline{Ae}(k) + \underline{bw}(k) \tag{28}$$

$$v(k) = \underline{h}^T\underline{e}(k) + dw(k) + p(k) \tag{29}$$

$$w(k) = \underline{\phi}^T(k)\underline{x}(k) - \alpha\underline{x}^T(k)\underline{\Gamma x}(k)v(k) \tag{30}$$

$$\underline{\phi}(k+1) = \underline{\phi}(k) - \underline{\Gamma}v(k)\underline{x}(k), \tag{31}$$

where $\underline{\Gamma} = \underline{\Gamma}^T > 0$ and $\alpha > 1/2$, consider

$$V(k) = \underline{e}^T(k)\underline{Pe}(k) + \tfrac{1}{2}\underline{\phi}^T(k)\underline{\Gamma}^{-1}\underline{\phi}(k), \tag{32}$$

where, due to the required [1]-[3] strict positive reality(SPR) of ($\underline{A}$, $\underline{b}$, $\underline{h}$, d), from the Kalman-Yacubovich lemma

$$\underline{A}^T\underline{PA} - \underline{P} = -\underline{qq}^T - \varepsilon\underline{L} \tag{33}$$

$$\underline{A}^T\underline{Pb} = \tfrac{1}{2}\underline{h} + \nu\underline{q} \tag{34}$$

$$\underline{b}^T\underline{Pb} = d - \nu^2 \tag{35}$$

where $\underline{L} = \underline{L}^T > 0$ and $\varepsilon, \nu > 0$. Evaluating

$$\Delta V(k) \triangleq V(k+1) - V(k) \tag{36}$$

yields

$$\Delta V(k) = -[\underline{e}^T(k)\underline{q} - \nu w(k)]^2 - \varepsilon\underline{e}^T(k)\underline{Le}(k) - (\alpha - \tfrac{1}{2})v^2(k)\underline{x}^T(k)\underline{\Gamma x}(k) - p(k)w(k). \tag{37}$$

Due to the positive definiteness of ε, $\underline{\Gamma}$, and $\underline{L}$ and the requirement that $\alpha > \frac{1}{2}$, $V(k) \leq 0$ when $p(k) = 0$.

To continue evaluation of $\Delta V(k)$ for nonzero p(k), substitute (29) into (30) and solve for w(k) as

$$w(k) = [1 + \alpha d\underline{x}^T(k)\underline{\Gamma x}(k)]^{-1}[\underline{\phi}^T(k)\underline{x}(k) - \alpha\underline{x}^T(k)\underline{\Gamma x}(k) \cdot (\underline{h}^T\underline{e}(k) + p(k))]. \tag{38}$$

Substitution of (38) into (37) yields

$$\Delta V(k) = -[\underline{e}^T(k)\underline{q} - \frac{\nu\underline{\phi}^T(k)\underline{x}(k)}{1 + \alpha d\underline{x}^T(k)\underline{\Gamma x}(k)} + \frac{\nu\alpha\underline{x}^T(k)\underline{\Gamma x}(k)}{1+\alpha d\underline{x}^T(k)\underline{\Gamma x}(k)}(\underline{h}^T\underline{e}(k) + p(k))]^2 - \varepsilon\underline{e}^T(k)\underline{Le}(k)$$

$$- (\alpha - \tfrac{1}{2})v^2(k)\underline{x}^T(k)\Gamma\underline{x}(k) - \frac{\underline{\phi}^T(k)\underline{x}(k)p(k)}{1 + \alpha d\underline{x}^T(k)\Gamma\underline{x}(k)}$$

$$+ \frac{\alpha\underline{x}^T(k)\Gamma\underline{x}(k)}{1+\alpha d\underline{x}^T(k)\Gamma\underline{x}(k)} (\underline{h}^T\underline{e}(k) + p(k))p(k). \quad (39)$$

Observation 1: Assuming the p(k) remains bounded, i.e.

$$0 < |p(k)| < N < \infty \quad \forall k, \quad (40)$$

and that large $\|\underline{\phi}(k)\|$ implies large $\|\underline{x}^T(k)\underline{\phi}(k)\|$, i.e. sufficient excitation occurs, there exists some $0 < \beta < \infty$, independent of $\underline{x}(k)$, such that

$$\|\underline{e}(k)\|^2 + \|\underline{\phi}(k)\|^2 > \beta^2 \quad \forall k \quad (41)$$

for which $\Delta V(k) \leq 0$.

Justification of Observation 1: Three possibilities must be considered: (i) $\|\underline{e}\|>1$ and large relative to $\|\underline{\phi}\|$ (and $\|\underline{\phi}^T\underline{x}\|$) and $|p|$, (ii) $\|\underline{\phi}\|$ (and $\|\underline{\phi}^T\underline{x}\|$) > 1 and large relative to $\|\underline{e}\|$ and $|p|$, and (iii) both $\|\underline{e}\|$ and $\|\underline{\phi}\|$ (and $\|\underline{\phi}^T\underline{x}\|$) are greater than one and large relative to $|p|$. In each case negative quadratics of the large terms will dominate any sum leaving (39) negative. □
Therefore, given the interpretations of $\underline{e}$ and $\underline{\phi}$ in (26) and (17), for output and parameter estimate errors of summed squared norm greater than β^2, the algorithm causes their decay since $\Delta V(k)$ is negative.

To establish that once (41) is dissatisfied the increase in $\|\underline{e}\| + \|\underline{\phi}\|$ is bounded form

$$\begin{bmatrix} \underline{e}(k+1) \\ \underline{\phi}(k+1) \end{bmatrix} = \underline{F}(k) \begin{bmatrix} \underline{e}(k) \\ \underline{\phi}(k) \end{bmatrix} + \underline{G}(k)p(k). \quad (42)$$

Such a forced, time-varying system description is derived by solution of (29) and (30) for

$$v(k) = [1 + \alpha d\underline{x}^T(k)\Gamma\underline{x}(k)]^{-1}[\underline{h}^T\underline{e}(k) + d\underline{\phi}^T(k)\underline{x}(k) + p(k)], \quad (43)$$

its substitution into (31), and the substitution of (38) into (28) yielding

$$\underline{F}(k) = \begin{bmatrix} \underline{A} - \dfrac{\alpha\underline{x}^T(k)\Gamma\underline{x}(k)\underline{b}\underline{h}^T}{1 + \alpha d\underline{x}^T(k)\Gamma\underline{x}(k)} & \dfrac{\underline{b}\underline{x}^T(k)}{1 + d\underline{x}^T(k)\Gamma\underline{x}(k)} \\ \dfrac{-\Gamma\underline{x}(k)\underline{h}^T}{1 + \alpha d\underline{x}^T(k)\Gamma\underline{x}(k)} & \underline{I} - \dfrac{d\Gamma\underline{x}(k)\underline{x}^T(k)}{1 + \alpha d\underline{x}^T(k)\Gamma\underline{x}(k)} \end{bmatrix} \quad (44)$$

and

$$\underline{G}(k) = \begin{bmatrix} \dfrac{-\alpha\underline{b}\underline{x}^T(k)\Gamma\underline{x}(k)}{1 + \alpha d\underline{x}^T(k)\Gamma\underline{x}(k)} \\ \dfrac{-\Gamma\underline{x}(k)}{1 + \alpha d\underline{x}^T(k)\Gamma\underline{x}(k)} \end{bmatrix}. \quad (45)$$

Formation of (42) validates the approach proposed in [6] for interpretation of the reduced-order modelling problem (and in [2] for the noisy measurement problem) with the algorithms of [1]-[3] such that the unmodeled output serves as a forcing function into the homogeneous equations of the adequately ordered case. Due to the boundedness of $\underline{A}$, $\underline{b}$, $\underline{h}$, and d and the normalization of $\underline{x}(k)$, $\underline{F}(k)$ and $\underline{G}(k)$ are bounded, i.e.

$$0 < \|\underline{F}(k)\| < \gamma_1 < \infty \quad (46)$$

$$0 < \|\underline{G}(k)\| < \gamma_2 < \infty . \quad (47)$$

Observation 2: If p(k) is bounded, as in (40), and $\|\underline{e}(k)\| + \|\underline{\phi}(k)\|$ is bounded, as in the converse of (41), then the succeeding state norm $\|\underline{e}(k+1)\| + \|\underline{\phi}(k+1)\|$ is bounded.

Justification of Observation 2: Reformulate (42) using properties of matrix norms [7] as

$$\left\| \begin{matrix} \underline{e}(k+1) \\ \underline{\phi}(k+1) \end{matrix} \right\| \leq \|\underline{F}(k)\| \left\| \begin{matrix} \underline{e}(k) \\ \underline{\phi}(k) \end{matrix} \right\| + \|\underline{G}(k)\| \; |p(k)| \quad (48)$$

Given (40), (46), (47), and dissatisfaction of (41) such that

$$\left\| \begin{matrix} \underline{e}(k) \\ \underline{\phi}(k) \end{matrix} \right\|^2 = \|\underline{e}(k)\|^2 + \|\underline{\phi}(k)\|^2 \leq \beta^2 \quad (49)$$

(48) becomes

$$\left\| \begin{matrix} \underline{e}(k+1) \\ \underline{\phi}(k+1) \end{matrix} \right\| < \gamma_1\beta + \gamma_2 N . \quad \square \quad (50)$$

Combining Observations 1 and 2 leads to

Observation 3: If p(k) is bounded, as in (40), and large $\|\underline{\phi}(k)\|$ implies large $\|\underline{x}^T(k)\underline{\phi}(k)\|$, i.e. sufficient excitation occurs, then $\|\underline{e}(k)\| + \|\underline{\phi}(k)\|$ in (28)-(31) is asymptotically bounded and the objective in (23) is achieved.

Justification of Observation 3: If (41) is satisfied then from Observation 1 $\Delta V(k) \leq 0$ implies a decay in $\|\underline{e}(k+1)\|^2 + \|\underline{\phi}(k+1)\|^2$ until (41) is dissatisfied. Assume, without loss of generality, that

$$(\gamma_1\beta + \gamma_2 N)^2 > \beta^2 . \quad (51)$$

Then from Observation 2, once any initially bounded $\|\underline{e}(0)\|^2 + \|\underline{\phi}(0)\|^2$ decays inside β^2 it will never exceed $(\gamma_1\beta + \gamma_2 N)^2$. □

IV. SUFFICIENT EXCITATION REQUIREMENT

The validity of the insights of the last section rests on the satisfaction of the condition embodied in Observation 1 that for nonzero p(k) large $\|\underline{\phi}(k)\|$ implies large $\|\underline{x}^T(k)\underline{\phi}(k)\|$. This is similar to the nonorthogonality sufficient excitation condition in equation error estimation [8] required for consistent parameter estimation. However the $\underline{x}$ in (18) includes both estimator inputs and outputs unlike equation error.

The need for this sufficient excitation requirement on the error system in (28)-(31) is possibly best illustrated by composition of a forced feedback loop from (28)-(31) with state vector $[\underline{e}^T(k)\ \underline{\phi}^T(k)]^T$. Consider a nonzero forcing p(k) causing a large $\underline{\phi}(k)$. However, if $\underline{x}^T(k)\underline{\phi}(k)$ is very small then the energy in the time-varying component of (30)-(31) is not appreciably dissipated through the linear, time-invariant, SPR component of (28)-(29). Assume, for example, that $\underline{x}^T(k)\underline{\phi}(k) = 0$ for $k > 0$. From (31) $||\underline{\phi}||$ could clearly become incessantly larger. Such a possibility is noted in

Observation 4: In the perturbed error system of (28)-(31) the possibility exists that though p and $\underline{e}$ remain bounded, if $||\underline{\phi}^T\underline{x}||$ is small despite large $||\underline{\phi}||$, $||\underline{\phi}||$ can increase without bound.

Justification of Observation 4: Choose $\underline{x}$ such that:

$$\underline{x}^T(k)\underline{\phi}(k) = 0 , \qquad (52)$$

$$||\underline{x}(k)|| = 1 , \qquad (53)$$

and for the ith component of (31)

$$\mathrm{sgn}\{\phi_i(k)\} = -\ \mathrm{sgn}\{\Gamma_i x_i(k)v(k)\} , \qquad (54)$$

where p(k) is chosen such that $v(k) \neq 0$, for all $k > 0$. Since α is unity in (19), for Γ chosen as the identity matrix from (53)

$$-\ \alpha\underline{x}^T(k)\Gamma\underline{x}(k) = -1 . \qquad (55)$$

In (30) the connection of $\underline{\phi}(k)$ through $\underline{x}^T(k)$ to the summation yielding w(k) can be broken due to (52) and

$$\frac{V(z)}{P(z)} = [1+d+\underline{h}^T(z\underline{I}-\underline{A})^{-1}\underline{b}]^{-1} . \qquad (56)$$

Since $d+\underline{h}^T(z\underline{I}-\underline{A})^{-1}\underline{b}$ is SPR, the hyperstability theorem [1] assures that, for any constant negative feedback $-\zeta$ around $d+\underline{h}^T(z\underline{I}-\underline{A})^{-1}\underline{b}$,

$$G(z) = \frac{d+\underline{h}^T(z\underline{I}-\underline{A})^{-1}\underline{b}}{1+\zeta[d+\underline{h}^T(z\underline{I}-\underline{A})^{-1}\underline{b}]} \qquad (57)$$

is stable. Since (56) and (57) share the same denominator for $\zeta = 1$, $\frac{V(z)}{P(z)}$ is stable and, for a bounded perturbation p, the output v and state $\underline{e}$ remain bounded assuming no unstable cancellations in (56) or $d+\underline{h}^T(z\underline{I}-\underline{A})^{-1}\underline{b}$. However from (31) and (54) ϕ_i (at least) can be seen to increase indefinitely such that $||\underline{\phi}||$ is unbounded. Such cases have been simulated. □

This does not imply possible instability of the AIF in (1)-(5) in "matching" (21) since the bounding normalization of $\underline{x}$ in (53) and the sign constraint in (54) are not generally achievable due to (18). The question still remains if for a bounded u and y_U (or p) the $\underline{\phi}$ of (17) can grow without bound due to the boundedness of $\underline{\phi}^T\underline{x}$. Observation 4 simply indicates that the answer to this question cannot be derived from examination of (28)-(31) without significant constraints on $\underline{x}$.

V. CONCLUSIONS

This paper established the bounded-input, bounded-output (BIBO) character of the error system associated with reduced-order use of HARF and similar adaptive recursive filters and output error identifiers given a sufficient excitation condition. This BIBO property supports the conclusions that, since a stable low-order model of a stable higher-order system always exists, the output estimates and parameter estimates of an AIF such as HARF will remain bounded despite reduced-order application given sufficient excitation. This sufficient excitation condition is peculiar in that it excludes the near-orthogonality of the vector of autoregressive and moving-average parameter estimate errors with the vector of past outputs and inputs of the AIF when the parameter error vector is large. Whether or not this undesirable condition can occur for an AIF such as HARF has not been ascertained. It was shown that it is possible for the general underlying perturbed error system derived from [4]. This BIBO robustness in reduced-order application can also be deduced [9] from the exponential convergence character of HARF in adequate-order application under a sufficient excitation condition on the filter input alone [10].

REFERENCES

[1] C. R. Johnson, Jr., "A convergence proof for a hyperstable adaptive recursive filter," IEEE Trans. on Info. Thy., vol. IT-25, no. 6, pp. 745-749, November 1979.

[2] I. D. Landau, "Unbiased recursive identification using model reference adaptive techniques," IEEE Trans. on Aut. Control, vol. AC-21, no. 2, pp. 194-202, April 1976.

[3] I. D. Landau, "An addendum to 'Unbiased recursive identification using model reference adaptive techniques'," IEEE Trans. on Aut. Control, vol. AC-23, no. 1, pp. 97-99, Feb. 1978.

[4] Y.-H. Lin and K. S. Narendra, "A new error model for adaptive systems," IEEE Trans. on Aut. Control, vol. AC-25, no. 3, pp. 585-587, June 1980.

[5] C. R. Johnson, Jr., "Another use of the Lin-Narendra error model: HARF," IEEE Trans. on Aut. Control, vol. AC-25, no. 5, pp. 985-988, October 1980.

[6] R. R. Bitmead and B. D. O. Anderson, "Performance of adaptive estimation algorithms in dependent random environments," IEEE Trans. on Aut. Control, vol. AC-25, no. 4, pp. 788-794, August 1980.

[7] H. H. Rosenbrock and C. Storey, Mathematics of Dynamical Systems. New York: John Wiley & Sons, p. 118, 1970.

[8] C. R. Johnson, Jr., "Adaptive parameter matrix and output vector estimation via an equation error formulation," IEEE Trans. on Syst., Man and Cybernetics, vol. SMC-9, no. 7, pp. 392-397, July 1979.

[9] B. D. O. Anderson and C. R. Johnson, Jr., "On reduced order output error identification and adaptive IIR filtering," VPI&SU Dept. of Elec. Eng. Tech. Report, May 1980.

[10] B. D. O. Anderson and C. R. Johnson, Jr., "Exponential convergence of adaptive identification and control algorithms," Univ. of Newcastle Dept. of Elec. Eng. Tech. Report, April 1980.