on Decision and Control
Honolulu, Hawaii • December 1990

Parameter Drift Instability in Adaptive Systems

DALE A. LAWRENCE
Department of Electrical and Computer Engineering
University of Cincinnati
Cincinnati, OH 45221

WILLIAM A. SETHARES
Department of Electrical and Computer Engineering
University of Wisconsin
Madison, WI 53706

WEI REN
Coordinated Science Laboratory
University of Illinois
Urbana, IL 61801

ABSTRACT

An error model describing a variety of adaptive systems is studied, representing applications in adaptive filtering, recursive parameter identification, and model reference adaptive control. For the *a posteriori* or normalized type of adaptive law in the "output error form", it is shown that error model instability requires a specific type of drift instability. Such drift is "non-explosive", but can lead to unbounded parameter and prediction errors. Drift is known to occur in the presence of algorithm disturbances, but a recent conjecture suggests it cannot happen in the ideal, no-disturbance case. Concrete examples are given showing drift to occur in this ideal case, provided an SPR condition is violated.

INTRODUCTION

Adaptive systems operating under disturbance-free conditions are exponentially asymptotically stable to their tuned parameterizations, provided that an internal strictly positive real (SPR) condition is satisfied, and certain signals are persistently exciting (PE). This strong form of stability provides a degree of robustness to non-zero algorithm disturbances. However, in applications such as output error identification [1,2] and model reference adaptive control [3–5], the SPR and PE conditions can be violated, making it reasonable to ask what stability properties are retained in their absence.

As shown in the well-known "counterexample to adaptive control" [6], stability of model reference adaptive control could fail catastrophically when the algorithm was operated in the presence of arbitrarily small sinusoidal disturbances at certain frequencies. The resulting "bursts" were analyzed in [8] and were attributed to the lack of PE, which allows disturbances to push parameter estimates into regions of unstable closed loop behavior. Similar examples for identification and adaptive filtering have appeared in [9–11]. More recently, the bursting mode of failure has been examined [12-15], showing that chaotic dynamics may be present.

In some cases [16], the adaptive algorithm recovers from bursts, so that parameter and prediction errors remain bounded. Heuristically, this has been argued as due to self generated excitation caused by the burst, which supplies PE and a resulting "cap" on the parameter estimates and errors. However, in other cases [11], it is shown that disturbances can conspire with decaying excitation to "mask" parameter motion, so that arbitrarily large parameter and prediction errors occur.

The self-stabilization effect of bursting has been examined in the case where the SPR condition is also violated. In this case, the tuned parameterization can be exponentially unstable, yet the algorithm can still be stable, in the sense that parameter and prediction errors are bounded. The results of [17,18] demonstrate algorithm input-output and Lyapunov stability under a sufficient power (SP) condition on the information vector. Even in the absence of the SP condition, simulations suggest that the algorithm parameters and errors remain bounded [19,25]. This has been argued using local stability results based on averaging theory. Away from the region of local instability, there exists a strong contractive tendency in the algorithm due to large signal sizes, resulting in overall limit cycle behavior.

On the other hand, more recent arguments predict that self-stabilization may not occur in the non-SPR case [14,15,20]. Using a scalar adaptive algorithm in the presence of PE and disturbances, an unstable manifold is shown to exist, along which parameter estimates may diverge. Although the non-PE, non-disturbance case is not (and probably cannot) be explicitly addressed, it can be viewed in terms of a "limit" as one of the bifurcation parameters approaches zero. The implication is that parameter drift to infinity should be possible in the disturbance free, non-PE, non-SPR case.

The question remains: are parameter estimates and prediction errors bounded in the "ideal" case (no disturbances), even if the SPR and PE conditions are violated? The present paper settles this question by exhibiting concrete examples of parameter drift in the non-SPR, non-PE, non-disturbance case. These examples concern the output error identification algorithm in the normalized, or *a posteriori* form. It is also shown that for certain *a posteriori* form algorithms, unbounded parameter and prediction errors can only occur if this type of drift instability is present. The implication of these results is that understanding and preventing drift is the key to maintaining stable adaptive system behavior.

I. ROLE OF DRIFT IN ADAPTIVE SYSTEM STABILITY

Simulation evidence [16,19,25] has indicated that certain normalized, or *a posteriori* algorithms may not possess an "explosive" form of instability, since large signals tend to eventually cause a kind of contraction, or re-stabilization effect. This section formalizes this notion of "non-explosive" instability, and shows that unbounded prediction and parameter errors can only occur if a particular kind of "drift" instability is present in the algorithm.

The class of adaptive systems considered are given by the following error model description

$$e_{k+1} = G(q^{-1})(v_{k+1}) - w_{k+1} = \phi_k^T \tilde{\theta}_{k+1} \tag{1}$$

$$\tilde{\theta}_{k+1} = \tilde{\theta}_k - h\phi_k v_{k+1} \tag{2}$$

where

e_{k+1} is the (*a posteriori*) equation error.

$G(q^{-1})$ is a monic n^{th}-order polynomial in the backward shift operator q^{-1}.

v_{k+1} is the measured (*a posteriori*) prediction error.

w_{k+1} is a disturbance.

ϕ_k is the information vector.

$\tilde{\theta}_k$ is the parameter error vector.

h is the step size.

This type of error system occurs in a variety of adaptive systems. Two common examples are explicit model reference adaptive control [3,4] and adaptive IIR filtering [1,2]. For implementation, a normalized *a priori* form of this algorithm is used (see e.g. [2]).

Equations (1) and (2) form a feedback system (the error system), which is time varying, and in general, nonlinear. A block diagram of this error system is shown in Figure 1.

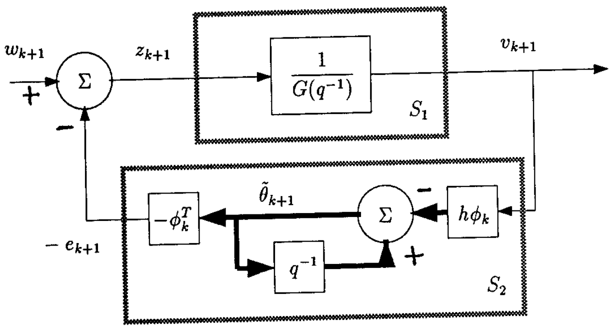

Figure 1: Error system model of the adaptive algorithms considered.

Our analysis of the role of drift in the stability of this error system is based on the π-sharing or dissipative systems viewpoint of [18]. This requires a state space representation of the individual subsystems S_1 and S_2 in Figure 1. A state space representation of S_2 is easily found from (1) and (2):

$$\tilde{\theta}_{k+1} = A_2\tilde{\theta}_k + B_{2,k}v_{k+1} \quad (3)$$
$$-e_{k+1} = C_{2,k}^T\tilde{\theta}_k + d_{2,k}v_{k+1} \quad (4)$$

where

$$A_2 = I \qquad B_{2,k} = -h\phi_k \quad (5)$$
$$C_{2,k} = -\phi_k \qquad d_{2,k} = h\phi_k^T\phi_k \quad (6)$$

For S_1, the following state space representation will be convenient

$$X_{k+1} = A_1X_k + B_1z_{k+1} \quad (7)$$
$$v_{k+1} = C_1^TX_k + d_1z_{k+1} \quad (8)$$

where with $G(q^{-1}) = 1 - \sum_{i=1}^n g_iq^{-i}$

$$A_1 = \begin{bmatrix} g_1 & \cdots & g_{n-1} & g_n \\ 1 & & 0 & 0 \\ & \ddots & & \vdots \\ 0 & & 1 & 0 \end{bmatrix} \qquad B_1 = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \quad (9)$$

$$C_1^T = \begin{bmatrix} g_1 & \cdots & g_{n-1} & g_n \end{bmatrix} \qquad d_1 = \begin{bmatrix} 1 \end{bmatrix}. \quad (10)$$

The following result can then be obtained.

LEMMA For the system defined by (1) and (2), if the disturbance sequence w is bounded (e.g. by $\|w\|_\infty$), and if there exists a positive β such that for some N, $\phi_k^T\phi_k \geq \beta X_k^TX_k$ whenever $X_k^TX_k \geq N$, then for all fixed step sizes $h > 0$, there exists an M such that $v_{k+1}^2 \geq M$ implies

(a) $\tilde{\theta}_{k+1}^T\tilde{\theta}_{k+1} \leq \tilde{\theta}_k^T\tilde{\theta}_k$ (parameter errors are monotone decreasing in norm), and

(b) $|v_{k+1}| \leq 2\|w\|_\infty + \sqrt{\tilde{\theta}_k^T\tilde{\theta}_k/h}$ (prediction errors are bounded in terms of $|\tilde{\theta}|$).

Proof: See Appendix.

Remark: The relative growth condition between the vectors X and ϕ (given by the bound β in the Lemma) is automatically satisfied in parallel and series-parallel identification structures. In the series-parallel case (e.g. equation error identification), $G(q^{-1}) = 1$, hence the system S_1 of (7)-(8) has no dynamics and X_k can be taken as zero. In the parallel case (e.g. output error identification), ϕ contains past (bounded) inputs and predicted outputs, hence cannot grow faster than the vector X of prediction errors, assuming the plant output is also bounded. The Lemma is satisfied in both cases. In model reference adaptive control, ϕ contains past plant inputs and outputs, so the Lemma is always satisfied. In fact, ϕ may actually grow faster than X unless the plant is minimum phase, showing that the condition of the Lemma is complementary to the usual "linear boundedness" assumption as in [5,7].

The two conclusions (a) and (b) in the Lemma imply the following results.

i.) The prediction error v cannot diverge to ∞ (i.e. it cannot have a limit of ∞). (If it did, then there would exist a K such that $|v_{k+1}| > M$, $\forall k > K$, and by (a), $\tilde{\theta}$ is monotone decreasing for all $k > K$. But this contradicts (b), since $\tilde{\theta}_k^T\tilde{\theta}_k \leq \tilde{\theta}_K^T\tilde{\theta}_K$ for all $k > K$.)

ii.) $\tilde{\theta}_k$ can increase without bound only on intervals where v_{k+1}^2 is bounded by M. (Obvious from (a)).

iii.) v can be unbounded only if $\tilde{\theta}$ is unbounded. (Obvious from (b) if $v^2 > M$, and v is clearly bounded otherwise).

iv.) v can be unbounded only if periods of large v (where $\tilde{\theta}$ is non-increasing) alternate with periods where v is small (where $\tilde{\theta}$ can be increasing without bound).

Therefore an "explosive" type of instability, where the prediction error v diverges to ∞, is not possible for this class of adaptive systems. Growth in v and $\tilde{\theta}$ must occur in alternate phases, each growing only when the other is well behaved. Point iii.) shows that the prediction error can be unbounded only if $\tilde{\theta}$ is unbounded, which occurs only on intervals where the prediction error is small. This type of instability in $\tilde{\theta}$ has been termed ***drift*** [9-11], since it is slower than an exponential escape to ∞, and occurs while the prediction error is well behaved. The above results show that unboundedness of prediction and parameter errors is due solely to the presence of drift instabilities in the adaptive system. Alternatively, preventing drift is the key to guaranteeing bounded prediction and parameter errors in this class of adaptive algorithms.

As the following sections show, the presence of drift instability depends on the properties of the excitation in the information vector ϕ, the disturbance sequence w, and the passivity properties of the transfer function $1/G(q^{-1})$.

II. Stability and Drift for SPR $1/G(q^{-1})$

When $1/G(q^{-1})$ is SPR, the error system (1), (2) can be shown to be stable in the following senses:

- In the absence of disturbances ($w = 0$), the equilibrium $\tilde{\theta} = 0$ is Lyapunov stable, and the prediction error $v \to 0$ [5,21].
- If the disturbances are in l^2 (square summable), then so are the prediction errors v (l^2-stability). In addition, the parameter errors $\tilde{\theta}$ are bounded. [18,22].

These results imply that disturbances (not in l^2) are necessary for drift instability, hence for unbounded prediction and parameter errors to occur, when $1/G(q^{-1})$ is SPR.

Disturbances are also sufficient to cause parameter drift instability, as shown in [9,10] for equation error adaptive filters, and recently in [11] for the output error case. This instability depends on conditions on the information vector sequence. If the information vector excitation is persistently rich [23], or is decaying rapidly enough [11], drift is prevented. Parameters can drift along subspaces of the parameter space which are "insufficiently excited", provided the disturbance serves to mask the effects of growing parameter errors as seen in the prediction error.

III. Stability Results for Non-SPR $1/G(q^{-1})$

The results of the previous section show that disturbances to the error system are necessary for drift to occur when $1/G(q^{-1})$ is SPR. However, when $1/G(q^{-1})$ is not SPR, stability can still be guaranteed if appropriate conditions are imposed on the excitation in the information vector ϕ.

Consider the algorithm (1), (2) in the following form

$$\tilde{\theta}_{k+1} = \tilde{\theta}_k - h\phi_k \frac{1}{G(q^{-1})}\{\phi_k^T \tilde{\theta}_{k+1}\} + h\phi_k \frac{1}{G(q^{-1})}\{w_{k+1}\} \qquad (11)$$

Suppose the parameter estimates are changing slowly enough so that

$$\tilde{\theta}_{k+1} \approx \tilde{\theta}_k - h\phi_k \frac{1}{G(q^{-1})}\{\phi_k^T\}\tilde{\theta}_{k+1} + h\phi_k \frac{1}{G(q^{-1})}\{w_{k+1}\} \qquad (12)$$

then using the matrix inversion lemma,

$$\tilde{\theta}_{k+1} \approx \left[I - \frac{h\phi_k \frac{1}{G(q^{-1})}\{\phi_k^T\}}{1 + h\phi_k^T \frac{1}{G(q^{-1})}\{\phi_k\}} \right] \left(\tilde{\theta}_k + h\phi_k \frac{1}{G(q^{-1})}\{w_{k+1}\} \right). \qquad (13)$$

The averaging theory of [24] provides a formal basis for the stability of the recursion (13). When the matrix

$$\sum_{i=k}^{k+L} \phi_i \frac{1}{G(q^{-1})}\{\phi_i^T\} \qquad (14)$$

is uniformly positive definite for some L and all k, the "average" effect is a contraction in $\tilde{\theta}$. If this contraction is strong enough, *total stability* is obtained, so that $\tilde{\theta}$ is bounded even in the presence of bounded disturbances. The condition (14) requires that the dominant part of the persistent excitation lies in frequency ranges where the transfer function associated with the operator $1/G(q^{-1})$ has a positive real part.

IV. Drift Instability with Non-SPR $1/G(q^{-1})$

An averaging argument also provides some insight into the possibility of algorithm drift instability. If "on the average", the matrix $\phi_k \frac{1}{G(q^{-1})}\{\phi_k^T\}$ is negative semi-definite, and the scalar $1 + h\phi_k^T \frac{1}{G(q^{-1})}\{\phi_k\}$ is positive, then the recursion (13) for $\tilde{\theta}$ is not contractive. These conditions are satisfied for $\phi_k^T\phi_k$ "small enough", and the operator $1/G(q^{-1})$ "non-positive" (e.g. non-SPR). Thus, $\tilde{\theta}$ may be growing rather than decreasing, even when no disturbances w are present. However, if $\tilde{\theta}$ grows too rapidly (e.g. exponentially) then the equation error $e = \phi^T\tilde{\theta}$ and hence v may grow large enough to violate the bounds presented in Section I, causing $\tilde{\theta}$ to decrease. Therefore, for unbounded drift in $\tilde{\theta}$ to occur, $\tilde{\theta}$ must grow at a rate slow enough to retain a bounded v. This unbounded drift scenario is possible, as shown by the following two concrete examples.

EXAMPLE 1 Drift of a numerator parameter in the predictor.

Plant: $y_{k+1} = \sum_{i=1}^{5} a_i y_{k-i+1} + \theta u_k$

A priori Predictor: $\bar{y}_{k+1} = \sum_{i=1}^{5} a_i \hat{y}_{k-i+1} + \hat{\theta}_k u_k$

A posteriori Predictor: $\hat{y}_{k+1} = \sum_{i=1}^{5} a_i \hat{y}_{k-i+1} + \hat{\theta}_{k+1} u_k$

A Priori prediction error (no disturbances): $\bar{v}_{k+1} = y_{k+1} - \bar{y}_{k+1}$

A Posteriori prediction error: $v_{k+1} = y_{k+1} - \hat{y}_{k+1}$

Parameter estimate update: $\hat{\theta}_{k+1} = \hat{\theta}_k + \dfrac{hu_k\bar{v}_{k+1}}{1 + hu_k^2}$

This results in the error system of (1) and (2) with

$$\tilde{\theta}_{k+1} = \theta - \hat{\theta}_{k+1} = \tilde{\theta}_k - hu_k v_{k+1} \qquad (15)$$

and

$$v_{k+1} = \frac{1}{A(q^{-1})}\{\tilde{\theta}_{k+1}u_k\} = \frac{1}{G(q^{-1})}\{e_{k+1}\} \qquad (16)$$

where $A(q^{-1}) = 1 - \sum_{i=1}^{5} a_i q^{-i}$, and e and v are the *a posteriori* errors. For simplicity, choose $A(q^{-1}) = (q^{-1} + p)^5$ where $p^{-1} = -\tan\pi/5$. Thus, $1/A(q^{-1})$ has magnitude $m = (p^2+1)^{-5/2}$ and phase $-\pi$ at the frequency $\omega = \pi/2$ rad/sec. To construct a drift example, choose the input u_k to cause

$$e_{k+1} = \tilde{\theta}_{k+1}u_k = \sin(k\pi/2 + \pi/4) \qquad (17)$$

so that when initial conditions are properly chosen, the steady state solution is obtained:

$$v_{k+1} = -m\sin(k\pi/2 + \pi/4). \qquad (18)$$

Substituting for u_k and v_{k+1} in the parameter error update yields

$$\tilde{\theta}_{k+1} = \tilde{\theta}_k + \frac{hm\sin^2(k\pi/2 + \pi/4)}{\tilde{\theta}_{k+1}} = \tilde{\theta}_k + \frac{hm}{2\tilde{\theta}_{k+1}} \qquad (19)$$

The two solutions for $\tilde{\theta}_{k+1}$ given $\tilde{\theta}_k$ are found from

$$\tilde{\theta}_{k+1} = \frac{\tilde{\theta}_k}{2} \pm \frac{1}{2}\sqrt{\tilde{\theta}_k^2 + 2hm} \qquad (20)$$

Taking the positive solution at each k shows that $\tilde{\theta}$ is monotone increasing if $\tilde{\theta}_o > 0$. Also, a supposed bound B on $\tilde{\theta}$ would imply $\tilde{\theta}_{k+1} \geq \tilde{\theta}_k + hm/(2B)$ from (19), for all k, contradicting any bound B. Hence $\tilde{\theta}$ is unbounded. The expression for e_{k+1} can then be used to find the required input sequence u_k. Note $u_k \to 0$, and thus $y_k \to 0$. The predictor output $\hat{y}_{k+1} \to \sin(k\pi/2+\pi/4)$. Finally, by specifying initial conditions and the input sequence u_k, all other signals in the system are uniquely specified by the predictor and parameter estimate update equations. Hence, the drift of $\tilde{\theta}$ to ∞ satisfies the chosen solution of (19).

In Example 1, the poles of the predictor are fixed at stable locations. It is also possible to have drift such that a pole of the predictor is unbounded:

EXAMPLE 2 Drift of a denominator parameter in the predictor.

Plant: $y_{k+1} = \sum_{i=2}^{5} a_i y_{k-i+1} + \theta y_k + bu_k$

A priori Predictor: $\bar{y}_{k+1} = \sum_{i=2}^{5} a_i \hat{y}_{k-i+1} + \hat{\theta}_k \hat{y}_k + bu_k$

A posteriori Predictor: $\hat{y}_{k+1} = \sum_{i=2}^{5} a_i \hat{y}_{k-i+1} + \hat{\theta}_{k+1} \hat{y}_k + bu_k$

A Priori prediction error (no disturbances): $\bar{v}_{k+1} = y_{k+1} - \bar{y}_{k+1}$

A Posteriori prediction error: $v_{k+1} = y_{k+1} - \hat{y}_{k+1}$

Parameter estimate update: $\hat{\theta}_{k+1} = \hat{\theta}_k + \dfrac{h\hat{y}_k \bar{v}_{k+1}}{1 + h\hat{y}_k^2}$

This results in the error system of (1) and (2) with

$$\tilde{\theta}_{k+1} = \theta - \hat{\theta}_{k+1} = \tilde{\theta}_k - h\hat{y}_k v_{k+1} \tag{21}$$

and

$$v_{k+1} = \frac{1}{A(q^{-1})}\{\tilde{\theta}_{k+1}\hat{y}_k\} = \frac{1}{G(q^{-1})}\{e_{k+1}\} \tag{22}$$

where $A(q^{-1}) = 1 - \theta q^{-1} - \sum_{i=2}^{5} a_i q^{-i}$, and e and v are the *a posteriori* errors. Choose $A(q^{-1})$ as in Example 1 and the predictor output $\hat{y}_k$ to cause

$$e_{k+1} = \tilde{\theta}_{k+1}\hat{y}_k = \sin(k\pi/2 + \pi/4) \tag{23}$$

so that when initial conditions are properly chosen, the steady state solution is obtained:

$$v_{k+1} = -m\sin(k\pi/2 + \pi/4)\,. \tag{24}$$

Substituting for $\hat{y}_k$ and v_{k+1} in the parameter error update yields

$$\tilde{\theta}_{k+1} = \tilde{\theta}_k + \frac{hm}{2\tilde{\theta}_{k+1}}\,. \tag{25}$$

As argued for (19), a solution of (25) for $\tilde{\theta}$ monotone increasing exists and is unbounded if $\tilde{\theta}_o > 0$. From (23), this implies that $\hat{y}_k \to 0$. The plant output can be found from $y_{k+1} = v_{k+1} + \hat{y}_{k+1} \to -m\sin(k\pi/2 + \pi/4)$. The required input is given by $bu_k = A(q^{-1})y_{k+1} \to \sin(k\pi/2 + \pi/4)$. As before, specifying initial conditions and the u_k sequence yields unique solutions for the system signals, and $\tilde{\theta}$ drifts to infinity.

It is important to note (cf. Section I) that if u_k is not carefully chosen to "hide" the growth of $\tilde{\theta}$ seen through the prediction error v, then the resulting growth of ϕ will eventually stop $\tilde{\theta}$ from further growth, and drift cannot occur. However, a scenario where v is unbounded can be constructed by selecting u on a time interval to cause parameters to drift to large values, then choose u on a following interval to expose them, i.e. provide excitation in the subspace where drift occurs. Choosing u on subsequent intervals to cause larger and larger drift excursions produces larger and larger "bursts" in v during the subsequent sufficiently excited intervals.

This example is a direct counterexample to the conjecture of [25]. However, while Example 2 has the same drift behavior as Example 1 according to the analysis, large magnitudes of $\tilde{\theta}$ are not likely for Example 2 in simulation or in practice since the predictor is unstable. The predictor output $\hat{y}$ will remain small (or go to zero) only if the errors in computing the input u_k are vanishingly small. This explains why this behavior was not observed in the simulation studies in [16,19,25]. In this sense, u_k can be thought of as an "open loop control" for an unstable system [11].

CONCLUSION

This paper has shown that common *a posteriori* adaptive algorithms (those satisfying the Lemma in Section I) are inherently robust in the sense that parameter estimates and prediction errors cannot diverge to infinity (i.e have a limit of ∞), even when disturbances are present, and PE or SPR conditions are violated. However, a failure mode still exists where unbounded parameters and signals occur. This failure mode is solely due to parameter drift instability.

Drift instability can exist when inadequate levels of persistent excitation are present. Disturbances are necessary to cause drift when the internal SPR condition is satisfied. This paper has demonstrated that drift can also occur without disturbances when this SPR condition is violated.

In the first example given here, the drift of $\tilde{\theta}$ does not lead to an unstable predictor, in the sense that freezing $\hat{\theta}$ at some large value does not result in an unstable transfer function. Hence, some small errors in computing the u_k sequence, which both masks growth in $\tilde{\theta}$ while causing that growth, can be tolerated. This "robustness" of the drift behavior has been verified by simulation, where errors in simulation initial conditions result in the same long term drift behavior, in spite of a period of transient errors in the calculated u_k sequence.

However, when drift occurs such that a pole of the predictor moves outside the unit circle (Example 2), errors in computing u_k are not tolerated as well, because they excite the unstable modes of this predictor. Here, drift is not expected to occur in practical applications, because computation and measurement errors are always available to perturb the delicate "masking" effect of the special excitation sequences causing/allowing drift.

ACKNOWLEDGEMENTS

The authors are grateful for many helpful discussions with P. V. Kokotovic, S. Dasgupta, P. R. Kumar, and C. Rohrs.

APPENDIX

Proof of Lemma 1: For the error system of Figure 1, note that

$$w_{k+1}v_{k+1} = z_{k+1}v_{k+1} - e_{k+1}v_{k+1}\,. \tag{A - 1}$$

For the $z_{k+1}v_{k+1}$ term, using (7) and (8)

$$\begin{aligned} z_{k+1}v_{k+1} &= X_{k+1}^T\Gamma_1 X_{k+1} - X_k^T\Gamma_1 X_k + p_1 v_{k+1}^2 + r_1 z_{k+1}^2 \\ &\quad - \begin{bmatrix} X_k^T & z_{k+1} \end{bmatrix} M_1 \begin{bmatrix} X_k \\ z_{k+1} \end{bmatrix} \end{aligned} \tag{A - 2}$$

where

$$M_1 = \begin{bmatrix} A_1^T\Gamma_1 A_1 - \Gamma_1 + p_1 C_1 C_1^T & A_1^T\Gamma_1 B_1 - C_1(1/2 - p_1 d_1) \\ B_1^T\Gamma_1 A_1 - C_1^T(1/2 - p_1 d_1) & B_1^T\Gamma_1 B_1 - d_1 + r_1 + p_1 d_1^2 \end{bmatrix} \tag{A - 3}$$

can be verified by direct substitution (see also [18]). With the choices

$$\Gamma_1 = \begin{bmatrix} n\epsilon & & & 0 \\ & (n-1)\epsilon & & \\ & & \ddots & \\ 0 & & & \epsilon \end{bmatrix}, \quad p_1 = 1/2 - n\epsilon, \quad r_1 = 1/2 \qquad (A-4)$$

and the definitions (9),(10), and $\epsilon = \sum_{i=1}^{n} g_i^2 = C_1^T C_1$,

$$M_1 = \begin{bmatrix} \frac{1}{2}C_1C_1^T - (C_1^T C_1)I & \underline{0} \\ \underline{0} & 0 \end{bmatrix}, \qquad (A-5)$$

hence M_1 is negative semi-definite and

$$z_{k+1}v_{k+1} \geq X_{k+1}^T\Gamma_1 X_{k+1} - X_k^T\Gamma_1 X_k + (1/2 - n\epsilon)v_{k+1}^2 + (1/2)z_{k+1}^2 . \qquad (A-6)$$

The state vector X_k is made up of past v_{k+1} (compare (1) and (8)-(10)), yielding

$$\begin{aligned} z_{k+1}v_{k+1} \;\geq\; & n\epsilon v_{k+1}^2 + (n-1)\epsilon v_k^2 + \cdots + \epsilon v_{k-n+2}^2 \\ & -n\epsilon v_k^2 - \cdots - 2\epsilon v_{k-n+2}^2 - \epsilon v_{k-n+1}^2 \\ & + (1/2 - n\epsilon)v_{k+1}^2 + (1/2)z_{k+1}^2 \\ = \; & -\epsilon X_k^T X_k + (1/2)v_{k+1}^2 + (1/2)z_{k+1}^2 . \end{aligned} \qquad (A-7)$$

Similarly, for the $-e_{k+1}v_{k+1}$ term in (A–1),

$$\begin{aligned} -e_{k+1}v_{k+1} \;=\; & \tilde{\theta}_{k+1}^T\Gamma_2\tilde{\theta}_{k+1} - \tilde{\theta}_k^T\Gamma_2\tilde{\theta}_k + p_2 e_{k+1}^2 + r_{2,k}v_{k+1}^2 \\ & - \begin{bmatrix} \tilde{\theta}_k^T & v_{k+1} \end{bmatrix} M_{2,k} \begin{bmatrix} \tilde{\theta}_k \\ v_{k+1} \end{bmatrix} \end{aligned} \qquad (A-8)$$

where $M_{2,k}$ is given by

$$\begin{bmatrix} A_2^T\Gamma_2 A_2 - \Gamma_2 + p_2 C_{2,k}C_{2,k}^T & A_2^T\Gamma_2 B_{2,k} - C_{2,k}(1/2 - p_2 d_{2,k}) \\ B_{2,k}^T\Gamma_2 A_2 - C_{2,k}^T(1/2 - p_2 d_{2,k}) & B_{2,k}^T\Gamma_2 B_{2,k} - d_{2,k} + r_{2,k} + p_2 d_{2,k}^2 \end{bmatrix} \qquad (A-9)$$

With the choices

$$\Gamma_2 = \frac{1}{2h}I, \quad p_2 = 0, \quad r_{2,k} = \frac{h}{2}\phi_k^T\phi_k \qquad (A-10)$$

together with the definitions (5) and (6), $M_{2,k} = [0]$ results, hence

$$-e_{k+1}v_{k+1} = \frac{1}{2h}\left(\tilde{\theta}_{k+1}^T\tilde{\theta}_{k+1} - \tilde{\theta}_k^T\tilde{\theta}_k\right) + \frac{h}{2}\phi_k^T\phi_k v_{k+1}^2 . \qquad (A-11)$$

Combining (A–11) and (A–7) into (A–1), and neglecting the $(1/2)z_{k+1}^2$ term,

$$\frac{1}{2h}\left(\tilde{\theta}_{k+1}^T\tilde{\theta}_{k+1} - \tilde{\theta}_k^T\tilde{\theta}_k\right) \leq w_{k+1}v_{k+1} - \left(\frac{1}{2} + \frac{h}{2}\phi_k^T\phi_k\right)v_{k+1}^2 + \epsilon X_k^T X_k . \qquad (A-12)$$

Under the hypothesis of the Lemma that ϕ is large whenever X is, i.e. there exists a positive β such that for some N, $\phi_k^T\phi_k \geq \beta X_k^T X_k$ whenever $X_k^T X_k \geq N$,

$$\tilde{\theta}_{k+1}^T\tilde{\theta}_{k+1} - \tilde{\theta}_k^T\tilde{\theta}_k \leq 2h\left[w_{k+1}v_{k+1} - \frac{1}{2}v_{k+1}^2 + X_k^T X_k\left(\epsilon - \frac{\beta h}{2}v_{k+1}^2\right)\right] \qquad (A-13)$$

when $X_k^T X_k \geq N$. Thus $\tilde{\theta}_{k+1}^T\tilde{\theta}_{k+1} \leq \tilde{\theta}_k^T\tilde{\theta}_k$ whenever $|v_{k+1}| > \max\left(2\|w\|_\infty , \sqrt{2\epsilon/\beta h}\right)$. (Here, $\|w\|_\infty$ is the bound on the disturbance w).

On the other hand, when $X_k^T X_k < N$, ignoring the $\frac{h}{2}\phi_k^T\phi_k v_{k+1}^2$ term from (A—11) yields

$$\tilde{\theta}_{k+1}^T\tilde{\theta}_{k+1} - \tilde{\theta}_k^T\tilde{\theta}_k \leq 2h\left(w_{k+1}v_{k+1} - \frac{1}{2}v_{k+1}^2 + \epsilon N\right) . \qquad (A-14)$$

Here, $\tilde{\theta}_{k+1}^T\tilde{\theta}_{k+1} \leq \tilde{\theta}_k^T\tilde{\theta}_k$ whenever $|v_{k+1}| > 2\|w\|_\infty + \sqrt{2\epsilon N}$. Thus, there exists an M such that both (A–13) and (A–14) hold, and $\tilde{\theta}^T\tilde{\theta}$ is non-increasing for $v^2 > M$ as claimed in (a).

For part b.), use either (A–13) or (A–14) to obtain

$$-\tilde{\theta}_k^T\tilde{\theta}_k \leq 2h\left(w_{k+1}v_{k+1} - \frac{1}{2}v_{k+1}^2\right) \qquad (A-15)$$

for $v_{k+1}^2 \geq M$, irrespective of $X_k^T X_k$, which implies that

$$|v_{k+1}| \leq 2\|w\|_\infty + \sqrt{\tilde{\theta}_k^T\tilde{\theta}_k/h} \qquad (A-16)$$

as claimed in part (b).

REFERENCES

[1] C. R. Johnson, Jr., "Adaptive IIR Filtering: Current Results and Open Issues", *IEEE Transactions on Information Theory*, Vol. IT-30, no. 2, March, 1984, pp. 237–250.

[2] C. R. Johnson, Jr., *Lectures in Adaptive Parameter Estimation*, Prentice Hall, 1987.

[3] I. D. Landau, *Adaptive Control: The Model Reference Approach*, Marcel Dekker, 1979.

[4] I. D. Landau and R. Lozano, "Unification of Discrete Time Explicit Model Reference Adaptive Control Designs", *Automatica*, Vol. 17, no. 4, 1981, pp. 593–611.

[5] G. C. Goodwin and K. S. Sin, *Adaptive Filtering, Prediction, and Control*, Prentice Hall, 1984.

[6] C. E. Rohrs, L. Valavani, M. Athans, and G. Stein, Analytical Verification of Undesirable Properties of Direct Model Reference Adaptive Control Algorithms", *Proc. 20th IEEE Conf. on Decision and Control*, San Diego, CA, Dec., 1981, pp. 1272–1284.

[7] G. C. Goodwin, P. J. Ramadge, and P. E. Caines, "Discrete Time Multivariable Adaptive Control", *IEEE Transactions on Automatic Control*, Vol. AC-25, no. 3, June, 1980, pp. 449–456.

[8] B. D. O. Anderson, "Adaptive Systems, Lack of Persistency of Excitation, and Bursting Phenomena", *Automatica*, Vol. 21, no. 3, 1985, pp. 247–258.

[9] W. A. Sethares, D. A. Lawrence, C. R. Johnson, Jr., and R. R. Bitmead, "Parameter Drift in LMS Adaptive Filters", *IEEE Transactions on Acoustics, Speech, and Signal Processing*, Vol. ASSP-34, No. 4, August, 1986, pp. 868–879.

[10] W. A. Sethares, D. A. Lawrence, and C. R. Johnson, Jr., "Parameter Drift in Equation Error Identification", *Proc. IEEE Conference on Decision and Control*, Ft. Lauderdale, FL, Dec., 1985.

[11] W. A. Sethares and D. A. Lawrence, "Output Error Identification is Not Globally Stable in the 'Nonideal' Case", *Proc 27th IEEE Conf. on Decision and Control*, Austin, TX, Dec., 1988, pp. 1528–1529.

[12] O. Macchi and M. Jaidane-Saidane, "Adaptive IIR Filtering and Chaotic Dynamics: Application to Audiofrequency Coding", *IEEE Transactions on Circuits and Systems*, Vol. 36, no. 4, April, 1989.

[13] W. A. Sethares, C. R. Johnson, Jr., and C. Rohrs, "Bursting in Adaptive Hybrids", *IEEE Transactions on Communications*, Aug., 1989.

[14] G. J. Rey, R. R. Bitmead, and C. R. Johnson, Jr., "The Dynamics of Bursting in Simple Adaptive Feedback Systems with Leakage", submitted to *IEEE Transactions on Circuits and Systems.*

[15] W. A. Sethares and I. M. Y. Mareels, "Dynamics of an Adaptive Hybrid", submitted to *IEEE Transactions on Circuits and Systems.*

[16] M. Jaidane-Sadane and O. Macchi, "Quasi-periodic Self Stabilization of Adaptive ARMA Predictors", *Int. J. of Adaptive Control and Signal Processing*, Vol. 2, no. 1, March, 1988, pp. 1-32.

[17] M. Tomizuka, "Parallel MRAS without Compensation Block", *IEEE Transactions on Automatic Control*, Vol. AC-27, no. 2, April, 1982, pp. 505–506.

[18] D. A. Lawrence and C. R. Johnson, Jr., "Recursive Parameter Identification Algorithm Stability Analysis via Π-Sharing", *IEEE Transactions on Automatic Control*, Vol. AC-25, no. 1, Jan., 1986, pp. 16–24.

[19] D. A. Schoenwald and P. V. Kokotovic, "On a Boundedness Conjecture for Output Error Adaptive Systems", *Proc. 27th IEEE Conf. on Decision and Control*, Austin, TX, Dec., 1988, pp. 780–781.

[20] L. Praly, "Oscillating Behavior and Fixes in Adaptive Linear Control: A Worked Example", *Proc 1988 IFAC Workshop on Robust Adaptive Control*, Newcastle, Australia, Aug., 1988.

[21] I. D. Landau, "A Stability Theorem with Applications to Adaptive Control", *IEEE Transaction on Automatic Control*, Vol.AC-24, no. 2 April, 1979, pp. 305–312.

[22] R. L. Kosut and B. Friedlander, "Robust Adaptive Control: Conditions for Global Stability", *IEEE Transactions on Automatic Control*, Vol. AC-30, no. 7, July, 1985, pp. 610–624.

[23] B. D. O. Anderson and C. R. Johnson, Jr., "On Reduced Order Adaptive Output Error Identification and Adaptive IIR Filtering", *IEEE Transactions on Automatic Control*, Vol. AC-27, Aug., 1982, pp. 927–933.

[24] B. D. O. Anderson, R. R. Bitmead, C. R. Johnson, Jr., P. V. Kokotovic, R. L. Kosut, I. M. Y. Mareels, L. Praly, and B. D. Riedle, *Stability of Adaptive Systems: Passivity and Averaging Analysis*, MIT Press, 1986.

[25] D. A. Schoenwald and P. V. Kokotovic, "Boundedness Conjecture for an Output Error Adaptive Algorithm" *Int. J. Adaptive Control and Signal Processing*, Vol. 4, 1990, pp. 27–47.