EXPONENTIAL CONVERGENCE AND ROBUSTNESS OF PERSISTENTLY EXCITED RECURSIVE-LEAST-SQUARES-WITH-FORGETTING OUTPUT ERROR IDENTIFICATION

by

D. A. Lawrence*, C. Richard Johnson, Jr.*, and B. D. O. Anderson**

ABSTRACT

This note demonstrates the exponential convergence of the recursive-least-squares-with-forgetting (RLSF) type output error identifier via the exponential stability of an associated, yet more general, error model formulation. This general error model approach encompasses several similarly well-behaved variants of output and equation error identification schemes. A persistent excitation property imposed on plant inputs alone, related to the order of the assumed plant ARMA model, is shown to provide the robust, exponentially convergent identification useful for such applications as those involving model order underestimation and/or plant parameter time-variation. The nonlinear, time-varying, forced error model associated with RSLF output error identification is derived for such non-ideal situations.

1. Introduction

Building upon the basic information available on the properties of many parameter identification *algorithms applied in "ideal" situations* [1,2,3] where such assumptions as plant linearity, time-invariance, and known model order are involved, recent attention has been directed toward analysis of algorithms in consideration of some of their less-than-ideal practical applications. Initial results are available for principal variants of equation error [4,5,6] and output error [5] identification for the robustness of the algorithm in the presence of disturbances caused by its "non-ideal" application. Owing to the similarity in structure of equation error and output error identifiers, results obtained for the simpler equation error forms have been extended to some corresponding output error forms. The results presented here follow this course in that the exponential stability and associated robustness of the recursive-least-squares-with-forgetting(RLSF)† equation error identifier are shown to characterize the corresponding RLSF output error algorithm, as conjectured in [6]. At the core of the present development, however lies a more general formulation. It includes many typical equation error and output error identification schemes as special cases, as well as many mixtures and minor variations. Thus a viewpoint is presented here that suggests a broad category of well-behaved, robust identification schemes.

We begin by stating the RLSF output error algorithm and manipulating it to fit a general error model formulation. Next, an exponential stability result is presented for the general error model, together with a discussion of the persistent excitation conditions involved. This is the key result required to show the exponential convergence of output error RSLF. Finally, recuced order modeling and plant time variation are shown to cause disturbance inputs to and changes in the state transition matrix of the homogeneous, exponentially stable error model of the ideal application. Therefore identification based on the form of the general error model presented here can be shown to possess robust behavior for suitably small disturbances by invoking the results of [7].

2. RSLF Output Error Formulation

Identification of a linear, time-invariant plant modelable by a difference equation of the form

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

will be considered, where y_M and u are the modeled plant output and plant input, respectively, ant it is desired that a collection of parameter estimates $a_i(k)$ and $b_j(k)$ converge to the true parameters a_i and b_j. For output-error-structure algorithms, a parallel model

$$z(k) \triangleq \sum_{i=1}^{n} \hat{a}_i(k+1) z(k-i) + \sum_{j=1}^{m} \hat{b}_j(k+1) u(k-j) \qquad (2)$$

is constructed (given here in the a posteriori form). The output z(k) of the parallel model is compared with the plant output $y_M(k)$ in obtaining a smoothed prediction error v(k)

$$v(k) \triangleq y_M(k) - z(k) + \sum_{i=1}^{n} \beta_i [y_M(k-i) - z(k-i)], \qquad (3)$$

where the β_i are coefficients to be selected subject to particular constraints [8].

*D. A. Lawrence and C. R. Johnson, Jr., are with the School of Electrical engineering, Cornell University, Ithaca, NY 14853, U.S.A. They are supported by National Science Foundation Grant ECS-8119312.

**B. D. O. Anderson is with the Department of Systems Engineering, Australian National University, Canberra, ACT 2600, Australia.

†This RLS nomenclature is based on the form of the direction matrix P update in (7) and is not necessarily indicative of a squared error minimization behavior, expecially in non-ideal applications of the output error form of RLS. This label is cumbersome, though presently conventional. We feel that the development of a universal, compact, unambiguous nomenclature is sorely needed in the field of recursive parameter estimation.

The parameter estimate update considered here can be compactly described by defining a parameter estimate vector

$$\hat{\theta}(k) \triangleq [\hat{a}_1(k),\ldots,\hat{a}_n(k),\ \hat{b}_1(k),\ldots,\hat{b}_m(k)]^T \quad (4)$$

and output error information vector

$$\psi_o(k) \triangleq [z(k-1),\ldots,z(k-n),\ u(k-1),\ldots,u(k-m)]^T \quad (5)$$

and requiring

$$\hat{\theta}(k+1) = \hat{\theta}(k) + \frac{1}{\alpha} P(k)\psi_o(k)v(k) \quad (6)$$

where $\alpha > 0$ and with $\gamma \in (0,1)$

$$P^{-1}(k+1) \triangleq \gamma(k)P^{-1}(k) + \psi_o(k)\psi_o^T(k). \quad (7)$$

$P(k_o)$ symmetric positive definite. Although the implied matrix inversion in (7) for use of P in (6) would be circumvented by use of the matrix inversion lemma [9] in an implementation of the algorithm, (7) remains a more tractable expression for analysis and will be used here throughout. By defining

$$\theta \triangleq [a_1,\ldots,a_n,\ b_1,\ldots,b_m]^T \quad (8)$$

the desired result for (6) will be that, while remaining bounded, both $\|\theta - \hat{\theta}(k)\| \to 0$ and $|v(k)| \to 0$ (and therefore $|y_M(k) - z(k)| \to 0$) as $k \to \infty$ for parameter identification.

We now introduce some additional definitions that will permit a system theoretic approach to the characterization and analysis of this algorithm. Accordingly, with

$$\tilde{\theta}(k) \triangleq \theta - \hat{\theta}(k) \quad (9)$$

$$e(k) \triangleq \psi^T(k)\tilde{\theta}(k+1) \quad (10)$$

$$X(k) \triangleq [y_M(k-1) - z(k-1),\ldots,y_M(k-n) - z(k-n)]^T \quad (11)$$

we form a general error model [10,11]

$$X(k+1) = AX(k) + Be(k) \quad (12)$$

$$v(k) = CX(k) + de(k) \quad (13)$$

$$\tilde{\theta}(k+1) = \tilde{\theta}(k) - \frac{1}{\alpha} P(k)\psi(k)v(k) \quad (14)$$

$$e(k) = \psi^T(k)\tilde{\theta}(k+1) \quad (15)$$

$$P^{-1}(k+1) = \gamma(k)P^{-1}(k) + \delta(k)\psi(k)\psi^T(k) \quad (16)$$

where for output error RLSF

$$A = \begin{bmatrix} a_1 & a_2 & \cdots & a_{n-1} & a_n \\ 1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & 0 \end{bmatrix},\quad B = \begin{bmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \quad (17)$$

$$C = [a_1 + \beta_1,\ldots,a_n + \beta_n], \quad d = 1$$

and

$$\psi(k) \equiv \psi_o(k)\ ,\ \gamma(k) \equiv \gamma\ ,\ \delta(k) \equiv 1. \quad (18)$$

In reformulating (1)-(8) to obtain (12)-(16), we make a shift in viewpoint to allow application of various stability theory tools and insights. For this general error model, we establish the following key preliminary lemmas.

LEMMA 1: For the error model of (12)-(16) if

(A1) (Strict positive reality (SPR) condition) The transfer function given by $C(zI-A)^{-1}B + (d-\rho)$ is (discrete) SPR for some $\rho > 0$.

(A2) (Designer-selected coefficients)

$$0 < \alpha < 2\rho$$

$$0 < \gamma(k) \le 1 \quad , \quad \forall k$$

$$0 < \delta(k) \le 2\frac{\rho}{\alpha} + \varepsilon \quad , \quad \text{some } \varepsilon > 0 \ , \ \forall k$$

(A3) (Initial condition)
$P(k_o)$ is symmetric positive definite

then

(R1) ("Equation error" boundedness and convergence)

$$\lim_{k\to\infty} e(k) = 0$$

(R2) (Smoothed output error boundedness and convergence)

$$\lim_{k\to\infty} v(k) = 0$$

(R3) (Past prediction error, i.e. (A,B,C,d)-system state, boundedness and convergence)

$$\lim_{k\to\infty} \|X(k)\| = 0.$$

Proof: A Lyapunov stability theory approach appears in [12], while a hyperstability theory approach appears in [13]. ΔΔΔ

The conditions in Lemma 1 yield a stable error model, but asymptotic stability is not implied and parameter errors $\tilde{\theta}$ need not tend to zero. The next three Lemmas provide conditions for parameter convergence to zero in increasingly stronger senses.

LEMMA 2: For the error model of (12)-(16) if (A1)-(A3) hold from Lemma 1, and

(A4) (Information vector persistent excitation lower bound). There exists a $\sigma_L > 0$ and an s such that

$$\sum_{i=k-s}^{k} \delta(i)\psi(i)\psi^T(i) > \sigma_L I \ , \ \forall k > k_o$$

then

(R4) (Asymptotic stability)

$$\lim_{k\to\infty} \left\| \begin{matrix} \tilde{\theta}(k) \\ X(k) \end{matrix} \right\| = 0.$$

Proof: See [12] and [13]. ΔΔΔ

LEMMA 3: For the error model of (12)-(16) if (A1)-(A4) hold from Lemmas 1 and 2, and

(A5) (Forgetting factor)
There exist fixed γ_o and γ_1 such that

$$0 < \gamma_0 \leq \gamma(k) \leq \gamma_1 < 1 \ , \ \forall k$$

then

(R5) (Exponential stability (e.s.)). There exists an $\eta \in (0,1)$ such that $\forall k_1 > k_0$, there exists an $M(k_1) > 1$ such that

$$\left\| \begin{matrix} \tilde{\theta}(k) \\ X(k) \end{matrix} \right\| < M(k_1)\eta^{(k-k_1)} \left\| \begin{matrix} \tilde{\theta}(k_1) \\ X(k_1) \end{matrix} \right\| \ , \ \forall k > k_1 .$$

Proof: See Appendix. ΔΔΔ

LEMMA 4: For the error model of (12)-(16) if (A1)-(A5) hold from Lemmas 1-3, and

(A6) (Information vector persistent excitation upper bound). There exists a $\sigma_U > \sigma_L$, and an s such that

$$\sigma_U > \sum_{i=k-s}^{k} \delta(i)\psi(i)\psi^T(i) \ , \ \forall k > k_0$$

then

(R6) (Uniform exponential stability (u.e.s)). There exists an $\eta \in (0,1)$ and an $M > 1$ such that

$$\left\| \begin{matrix} \tilde{\theta}(k) \\ X(k) \end{matrix} \right\| < M \ \eta^{(k-k_1)} \left\| \begin{matrix} \tilde{\theta}(k_1) \\ X(k_1) \end{matrix} \right\| \ , \ \forall k \geq k_1 \geq k_0 .$$

Proof: See appendix. ΔΔΔ

Remarks:

(i) The results of Lemmas 2 and 3, while too weak to use in answering questions about robustness, serve to point out those conditions responsible for strengthening the stability results and may be of primary importance in other applications. In particular, Lemma 2 includes the recursive-least-squares-without-forgetting algorithms in both the equation error and output error structures.

(ii) From the proof (see Appendix), the exponential rate constant η depends directly on the "degree of forgetting", γ_1, and on some notion of "degree of SPR", embodied in p and Q, for the system (A,B,C,d). M depends additionally on the persistent excitation bound ratio σ_U/σ_L.

(iii) The inclusion of $\delta(k)$ in the persistent excitation conditions (A4) and (A6) may prove useful in certain applications, e.g. [14] where $\delta(k)$ is used as a normalizing factor on the $\psi(k)$ sequence.

By making the special choices in (17) and (18), output error RLSF is exponentially convergent under the assumptions of Lemma 3. The statement of the Lemma is general enough to allow flexibility in the construction of the sequences $\gamma(k)$ and $\delta(k)$, perhaps subject to additional constraints based on the behavior of the $\psi(k)$ sequence, to improve particular aspects of the algorithm's performance. For example, when $\delta(k) = 0$, and $\gamma(k) = 1$ in (16), the parameter update assumes a gradient-like structure since P(k) becomes fixed, resulting in a relatively slow parameter estimate adjustment which may be desirable should the persistent excitation of $\psi(k)$ "momentarily fail" i.e., for a particular s and σ, (A4,A6) fail to hold for some k. (Given the same persistent excitation condition as in (A4,A6), this gradient version of output (and equation) error identification is proven exponentially stable in [5].) Thus mixtures of gradient, least squares, and RLSF variants are possible and may prove desirable, i.e. exponentially convergent, in mixtures needed for certain applications.

Additional freedom is provided in the Lemmas due to the unspecified source of information in $\psi(k)$. The equation error variations are included in the general error model form, for example, by choosing $\psi(k) = \psi_e(k)$ where

$$\psi_e(k) = [y_M(k-1),\ldots,y_M(k-n),\ u(k-1),\ldots,u(k-m)]^T \qquad (20)$$

and although typical equation error forms have an identity operator from e(k) to v(k), SPR operators as represented by (12) and (13) are also possible from the Lemmas. (Assumption (A1)).

The exponential stability of the error model hinges on the persistent excitation of $\psi(k)$. In output error algorithms, $\psi(k) \equiv \psi_0(k)$ from (5) contains plant inputs u as well as outputs z of the non-linear time-varying parallel identifier system (2). This raises the question of what conditions on u(k) alone can guarantee this persistent excitation of $\psi_0(k)$, and as shown below, persistent excitation of $\psi_0(k)$ is provided by the externally verifiable richness of just the plant input sequence u(k). To demonstrate this conclusion we exploit the similarities between output error and equation error behavior provided by Lemma 1. Following [6], where exponential convergence of output error RLSF was conjectured, we state our main result.

THEOREM If for the general error model (12)-(16) the assumptions (A1), (A2), (A3), (A5) of Lemmas 1-3 are satisfied and

(A7) (Persistently exciting plant input) There exist $\sigma_2 > \sigma_1 > 0$ and a fixed s $> 2n + m - 2$ such that

$$\sigma_2 I > \sum_{i=k}^{k+s-n+1} \begin{bmatrix} u(i+n) \\ \vdots \\ u(i-m+1) \end{bmatrix} [u(i+n)\ldots u(i-m+1)] > \sigma_1 I \ , \quad \forall k > k_0 .$$

(A8) (Exact modeling assumption) The dimension n is minimal for the plant modeled by (1).

then

(R7) (Uniform exponential stability (u.e.s.)) There exists an $\eta_0 \in (0,1)$ and an $M_0 > 1$ such that

$$\left\| \begin{matrix} \tilde{\theta}(k) \\ X(k) \end{matrix} \right\| > m_0 \eta_0^{(k-k_1)} \left\| \begin{matrix} \tilde{\theta}(k_1) \\ X(k_1) \end{matrix} \right\| \ , \ \forall k \geq k_1 \geq k_0 .$$

Proof: (A7) and (A8) are sufficient to insure a persistent excitation condition similar to (A4) and (A6) for the equation error information vector $\psi_e(k)$ of (20). See [3]. By application of Lemma 4 with $\psi(k) \equiv \psi_e(k)$, this equation error variation of the output error algorithm is u.e.s. *Further, since* $\psi_o(k) = \psi_e(k) - [X^T(k), 0_m^T]^T$, *by* application of Lemma 1 for $\psi(k) = \psi_o(k)$ we have $\psi_o(k) \to \psi_e(k)$ as $k \to \infty$ from (R3) for any $\psi_o(k)$ sequence. Now from (12)-(16) we form a more compact description by defining the error model state $\Omega(k) = [\tilde{\theta}^T(k), X^T(k)]^T$ so that

$$\Omega_o(k+1) = F_o(k,\Omega_o,\theta)\Omega_o(k) \tag{21}$$

for the output error system and

$$\Omega_e(k+1) = F_e(k,\theta)\Omega_e(k) \tag{22}$$

for the equation error system, where it is readily verified that

$$F_o(k,\Omega_o,\theta) = \begin{bmatrix} I - \dfrac{P(k)\psi_o(k)\psi_o^T(k)}{\alpha+\psi_o^T(k)P(k)\psi_o(k)} & \dfrac{-P(k)\psi_o(k)C}{\alpha+\psi_o^T(k)P(k)\psi_o(k)} \\ \dfrac{\alpha B\psi_o^T(k)}{\alpha+\psi_o^T(k)P(k)\psi_o(k)} & A - \dfrac{B\psi_o^T(k)P(k)\psi_o(k)C}{\alpha+\psi_o^T(k)P(k)\psi_o(k)} \end{bmatrix} \tag{23}$$

and $F_e(k,\theta)$ is $F_o(k,\Omega_o,\theta)$ with $\psi_o(k)$ replaced by $\psi_e(k)$. (The fact that the equation error system is linear is indicated by the absence of Ω in the argument of F_e.) It can then be seen that

$$\Omega_o(k+1) = [F_e(k,\theta) + \tilde{F}(k,\Omega_o,\theta)]\Omega_o(k) \tag{24}$$

where $F_e(k,\theta)$ is a bounded, u.e.s. state matrix, and $\|\tilde{F}(k,\Omega_o,\theta)\| \to 0$ as $k \to \infty$. Thus $F_o(k,\Omega_o,\theta)$ is u.e.s. (see Appendix) and the desired uniform exponential stability result (R7) follows. ΔΔΔ

Remark:

(v) Although the existence of persistent excitation bound constants and associated exponential convergence rate constants are provided, their interrelation is not made explicit by the proof. It is therefore difficult to obtain an a priori desired convergence rate by choice of constants and input sequences, and further work in making these connections is necessary.

3. Reduced-Order and Time-Varying Identification - An Application for Robustness

We will now describe the changes to the homogeneous error model of (12)-(16) (or (21) and (23)) in the practical case of reduced-order modeling of a time-varying plant. Suppose an accurate plant model consists of a time-varying version of (1) along with an additional unmodeled component so that the plant output is $y(k)$ where

$$y(k) = y_M(k) + y_U(k) \tag{25}$$

$$y_M(k) = \sum_{i=1}^{n} a_i(k)y_M(k-i) + \sum_{j=1}^{m} b_j(k)u(k-j). \tag{26}$$

Note that $y_U(k)$ may represent noise or unmodeled plant components. Introducing

$$\theta(k) \triangleq [a_1(k),\ldots,a_n(k),b_1(k),\ldots,b_m(k)]^T \tag{27}$$

$$\tilde{\theta}(k) \triangleq \theta(k) - \hat{\theta}(k) \tag{28}$$

$$\Delta\theta(k+1) \triangleq \theta(k) - \theta(k+1) \tag{29}$$

$$\nu(k) \triangleq y_U(k) + \sum_{i=1}^{n} \hat{a}_i y_U(k-i) \tag{30}$$

along with $A(k)$ and $C(k)$ defined as in (17) but with the a_i replaced by $a_i(k)$, we form the corresponding error model. With $\Omega_o(k)$ defined as before as $[\tilde{\theta}^T(k), X^T(k)]^T$ and $F_o(k,\Omega_o,\theta(k))$ as in (23) but with A and C replaced by $A(k)$ and $C(k)$, a little algebra yields

$$\Omega_o(k+1) = F_o(k,\Omega_o,\theta(k))\Omega_o(k) + G(k,\Omega_o)\rho(k) \tag{31}$$

where $\rho(k) = [\Delta\theta^T(k+1), \nu(k)]^T$ and

$$G(k,\Omega_o) = \begin{bmatrix} -I & \dfrac{-P(k)\psi_o(k)}{\alpha+\psi_o^T(k)P(k)\psi_o(k)} \\ \underline{0} & \dfrac{-B\psi_o^T(k)P(k)\psi_o(k)}{\alpha+\psi_o^T(k)P(k)\psi_o(k)} \end{bmatrix} \tag{32}$$

The error model in (31) for the time-invariant, reduced order identification problem ($\Delta\theta(k) \equiv 0$) is bounded-input, bounded-state stable provided that the homogenous unperturbed system satisfies the conditions of the Theorem, and $\nu(k)$ is suitably small. The essential *requirements are that the homogenous system is* assymptotically stable with u.e.s. linearization and that $G(k,\Omega_o)$ is bounded [7]. The Theorem provides the conditions for the u.e.s stability criterion and from (32) $G(k,\Omega)$ is readily seen to be bounded. As in the Theorem, the bound on the perturbation for this result appears in the proof in a non-constructive manner and a "suitably small" unmodeled component $y_U(k)$ is difficult to characterize a priori without further assumptions regarding the source of y_U. See [5] and [6] for additional relevant remarks.

A similar result can be shown to hold for the full-order time-varying identification task ($\nu(k) \equiv 0$) [7], provided that the time variation is slow enough, i.e., that $\Delta\theta(k)$ is suitably small. Again the exponential stability of the unperturbed error model together with certain smoothness conditions on the changes introduced by the perturbing input $\Delta\theta$ are the essential requirements.

The case of simultaneous occurence of nonzero ν and $\Delta\theta$ can be argued as follows. By noting *that the homogeneous version of (31)*

$$\bar{\Omega}_o(k+1) = F_o(k,\bar{\Omega}_o,\theta(k))\bar{\Omega}_o(k) \tag{33}$$

for any fixed value of $\theta(k)$ is u.e.s. by the Theorem, the results of [7] provide that this homogenous system is u.e.s. if $\Delta\theta(k)$ is suitably small. By regarding (31) as a perturbation to (33), suitably small bounds on both $\Delta\theta(k)$ and $\nu(k)$ result in bounded-input bounded-state stability for (31) as argued previously for the case when $\Delta\theta(k) \equiv 0$.

5. Conclusion

The class of identification algorithms modelable by the general error model of (12)-(16) has been shown uniformly exponentially convergent given a strictly positive real condition on a system containing unknown plant coefficients and given a persistently exciting input sequence. This class of algorithms includes the output error RLSF algorithm. Since the exponential convergence rate affects the degree of robustness obtained in non-ideal situations, prescription of this rate by choices in algorithm structure and input sequence would be highly desirable in practical applications, and remains an important focus for futher work.

6. Appendix

Proof of Lemma 3:

From (A1) if $C(zI-A)^{-1}B + (d-p)$ is SPR, then there exist positive definite Γ and Q, and matrix L and scalar r [1] satisfying

$$A^T\Gamma A - \Gamma = -LL^T - Q$$
$$A^T\Gamma B = C^T - Lr \qquad \text{(a-1)}$$
$$B^T B = 2(d-p) - r^2$$

so that by choosing the Lyapunov function

$$w(\tilde{\theta},X,k) = \tilde{\theta}^T(k)P^{-1}(k)\tilde{\theta}(k) + \frac{1}{\alpha}X^T(k)\Gamma X(k) \qquad \text{(a-2)}$$

it can be seen that under (A1)-(A3) we have

$$\Delta w(\tilde{\theta},X,k+1) \overset{\Delta}{=} w(\tilde{\theta},X,k+1) - w(\tilde{\theta},X,k)$$
$$\leq (\gamma(k)-1)\tilde{\theta}^T(k)P^{-1}(k)\tilde{\theta}(k) - \frac{1}{\alpha}X^T(k)QX(k)$$
$$- (2\frac{p}{\alpha} - \delta(k))e^2(k) \leq 0 \ , \ \forall k. \qquad \text{(a-3)}$$

This change is overbounded by a uniformly negative definite function of e(k) and X(k). Since $w(\tilde{\theta},X,k_0)$ is finite, $w(\tilde{\theta},X,k)$ converges and $\Delta w(\tilde{\theta},X,k)$ converges to zero, yielding (R1)-(R3). Now for all k

$$w(\tilde{\theta},X,k+1) \leq \gamma(k)\tilde{\theta}^T(k)P^{-1}(k)\tilde{\theta}(k) + \frac{1}{\alpha}X^T(k)[\Gamma-Q]X(k) \qquad \text{(a-4)}$$

so by letting $\lambda_{max}(\Gamma)$ and $\lambda_{min}(Q)$ indicate the maximum eigenvalue of Γ and minimum eigenvalue of Q respectively,

$$\Gamma > \left(\frac{\lambda_{max}(\Gamma) - \lambda_{min}(Q)}{\lambda_{max}(\Gamma)}\right)\cdot\Gamma \geq \Gamma - Q. \qquad \text{(a-5)}$$

Then if (A5) holds, there exists an $\eta \in (0,1)$ as

$$\eta > \max\left[\gamma_1 \ , \ \frac{\lambda_{max}(\Gamma) - \lambda_{min}(Q)}{\lambda_{max}(\Gamma)}\right] \qquad \text{(a-6)}$$

so we have

$$w(\tilde{\theta},X,k+1) < \eta\, w(\tilde{\theta},X,k) \ , \ \forall k \qquad \text{(a-7)}$$

Under (A4), we obtain a lower bound for $P^{-1}(k)$ as follows. From (16), we have

$$\sum_{j=k-s}^{k} P^{-1}(j+1) = \sum_{j=k-s}^{k} \gamma(j)P^{-1}(j) + \sum_{j=k-s}^{k} \psi(j)\delta(j)\psi^T(j) \qquad \text{(a-8)}$$

whence

$$P^{-1}(k+1) + \sum_{j=k-s+1}^{k} [1-\gamma(j)]P^{-1}(j) \qquad \text{(a-9)}$$
$$= \gamma(k-s)P^{-1}(k-s) + \sum_{j=k-s}^{k} \delta(j)\psi(j)\psi^T(j) \geq \sigma I$$

with the last inequality following from (A4). Now (16) also implies that

$$P^{-1}(j) \leq \frac{1}{\gamma(j)}P^{-1}(j+1) \leq \frac{1}{\gamma(k)\ldots\gamma(j)}P^{-1}(k+1). \qquad \text{(a-10)}$$

Hence

$$\sigma I \leq \left[1 + \frac{1-\gamma(k)}{\gamma(k)} + \ldots + \frac{1-\gamma(k-s+1)}{\gamma(k)\ldots\gamma(k-s+1)}\right]P^{-1}(k+1)$$
$$\leq (1/\gamma_0^{\,s})P^{-1}(k+1) \qquad \text{(a-11)}$$

(which is a greater lower bound than that given in [6]). Using this we have from (a-2) and (a-7)

$$\left\|\begin{matrix}\tilde{\theta}(k)\\X(k)\end{matrix}\right\| \leq \frac{w(\tilde{\theta},X,k)}{\min(\gamma_0^{\,s}\cdot\sigma_L \ , \ \frac{1}{\alpha}\lambda_{min}(\Gamma))}$$
$$< \frac{\eta^{(k-k_1)}w(\tilde{\theta},X,k_1)}{\min(\gamma_0^{\,s}\cdot\sigma_L \ , \ \frac{1}{\alpha}\lambda_{min}(\Gamma))}$$
$$\leq \eta(k-k_1)\underbrace{\frac{\max(\lambda_{max}(P^{-1}(k_1)) \ , \ \frac{1}{\alpha}\lambda_{max}(\Gamma))}{\min(\gamma_0^{\,s}\cdot\sigma_L \ , \ \frac{1}{\alpha}\lambda_{min}(\Gamma))}}_{M(k_1)}\left\|\begin{matrix}\tilde{\theta}(k_1)\\X(k_1)\end{matrix}\right\| \qquad \text{(a-12)}$$

As claimed in (R5). $\Delta\Delta\Delta$

Proof of Lemma 4:

Following [6], assumption (A6) yields an upper bound for $P^{-1}(k)$: There exists an N such that

$$P^{-1}(k) \geq \frac{\sigma_U\cdot I}{1-\gamma_1^{(s+1)}} + N\gamma_1^{\,k} \ , \ \forall k. \qquad \text{(a-13)}$$

Then from (a-12), there exists an M such that

$$M(k_1) = \frac{\max(\lambda_{max}(P^{-1}(k_1)) \ , \ \frac{1}{\alpha}\lambda_{max}(\Gamma))}{\min(\gamma_0^{\,s}\cdot\sigma_L \ , \ \frac{1}{\alpha}\lambda_{min}(\Gamma))} \leq M \ , \ \forall k_1. \qquad \text{(a-14)}$$

and result (R6) follows. $\Delta\Delta\Delta$

For the extension of F_e exponential stability to F_o exponential stability in the Theorem:

If $\|\Omega_e(k+k_1)\| < M_e\eta_e^k\|\Omega_e(k_1)\|$, $\forall k > 0$, $\forall k_1 \geq k_0$ and $\|\tilde{F}(k,\Omega,e)\| \to 0$ as $k\to\infty$, for some $\eta_e \in (0,1)$, $M_e > 1$, then for any η,

$$\|\Omega_o(k+n)\| \leq [\|F_e(k,\theta)\dots F_e(k+n-1,\theta)\| + \|S(k+n,k)\|]\cdot\|\Omega_o(k)\| \quad \text{(a-15)}$$

where $\|S(k+n,k)\| \to 0$ as $k\to\infty$, so for any $\tilde{\eta} > 0$ there exists a $K(\tilde{\eta},n)$ such that

$$\|\Omega_o(k+n)\| < M_e(\eta_e+\tilde{\eta})^n \cdot \|\Omega_o(k)\| \;,\; k > K(\tilde{\eta},n). \quad \text{(a-16)}$$

By choosing $\tilde{\eta}$ and n so that $\eta_e + \tilde{\eta} = \bar{\eta} < 1$, and $M_e(\bar{\eta})^n < 1$, then for any i and any $1 \leq j < n$, from (a-16)

$$\|\Omega_o(k+in+j)\| < (M_e(\bar{\eta})^n)^i \, M_e(\bar{\eta})^j \|\Omega_o(k)\| \quad \text{(a-17)}$$

for $k > K^*$, $K^* = \max\limits_{1\leq j\leq n} K(\tilde{\eta},j)$. Now since $M_e(\bar{\eta})^n < 1$, $M_e(1/n)\bar{\eta} < 1$, and since $M_e > 1$, $M_e{}^{(1/n)} > M_e{}^{(1/j)}$. Then from (a-17)

$$(M_e(\bar{\eta})^n)^i M_e(\bar{\eta})^j = (M_e{}^{(1/n)}\bar{\eta})^{ni}(M_e{}^{(1/j)}\bar{\eta})^j$$

$$< (M_e{}^{(1/n)}\bar{\eta})^{ni+j} \triangleq \eta_o{}^{(ni+j)}. \quad \text{(a-18)}$$

Since the state transitions for $k < K^*$ are finite, and any k can be written as $ni+j$ for some i, some $j < n$, there exists an M_o such that

$$\|\Omega_o(k+k_1)\| < M_o \eta_o{}^{(k-k_1)} \|\Omega_o(k_1)\| \;,\; \forall k,k_1 > k_o \quad \text{(a-19)}$$

and the F_o system is u.e.s. as claimed.

Alternatively, we could argue [15] that by Lemma 1, $\psi_o(k) \to \psi_e(k)$, so there exists a K^* such that ψ_o is sufficiently exciting if ψ_e is, for $k > K^*$. By Lemma 4 with $\psi = \psi_o$, F_o is u.e.s. for $k > K^*$. Since K^* is finite and the state transitions for $k < K^*$ are finite, there exists an M_o such that (R7) holds as claimed, with $\eta_o = \eta_e$. The earlier proof, however more complicated, may hold more promise in finding bounds on K^* and M_o for particular applications.

7. References

[1] I. D. Landau and H. M. Silveira, IEEE Trans. on Auto. Control, vol. AC-24, No. 2, April, 1979, pp. 305-311.

[2] J. M. Mendel, Discrete Techniques of Parameter Estimation: The Equation Error Formulation. New York: Marcel Dekker, 1973.

[3] B. D. O. Anderson and C. R. Johnson, Jr., Automatica, Vol. 18, No. 1, 1982, pp. 1-13.

[4] P. A. Ioannou and P. V. Kokotovic, IEEE Trans. on Auto. Control. Vol. AC-27, No. 4, Aug., 1982, pp. 921-927.

[5] B. D. O. Anderson and C. R. Johnson, Jr., IEEE Trans. on Auto. Control. Vol. AC-27, No. 4, Aug., 1982, pp. 927-932.

[6] R. M. Johnstone, C. R. Johnson, Jr., R. R. Bitmead, and B. D. O. Anderson, Systems and Control Letters, Vol. 2, No. 2, Aug., 1982, pp. 77-81.

[7] B. D. O. Anderson and R. M. Johnstone, Proc. 20th IEEE Conf. on Dec. and Control, Vol. 2, pp. 510, 1981.

[8] I. D. Landau, IEEE Trans. on Auto. Control, Vol. AC-21, No. 2, April, 1976, pp. 194-202.

[9] B. Noble, Applied Linear Algebra, Prentice-Hall, Englewood Cliffs, NJ, 1969, pp. 147.

[10] Y-H. Lin and K. S. Narendra, IEEE Trans. on Auto. Control, Vol. AC-25, No. 3, June, 1980, pp. 585-587.

[11] C. R. Johnson, Jr., IEEE Trans. on Auto. Control, Vol. AC-25, No. 5, Oct., 1980, pp. 985-987.

[12] D. A. Lawrence, MS Thesis, Cornell Univ., Ithaca, NY, June, 1982.

[13] I. D. Landau, IEEE Trans. on Auto. Control, Vol. AC-25, Aug., 1980, pp. 814-817.

[14] R. Lozano, Automatica, Vol. 19, No. 1, Jan., 1983, pp. 95-97.

[15] R. McG. Johnstone, PhD Thesis, University of Newcastle, NSW, Australia, June, 1982, pp. 65-68.