Automatica, Vol. 24, No. 2, pp. 217–225, 1988
Printed in Great Britain.

0005–1098/88 $3.00 + 0.00
Pergamon Press plc

Identification of Physical Parameters in Structured Systems*

SOURA DASGUPTA,† BRIAN D. O. ANDERSON‡ and R. JOHN KAYE§

Adaptive identification algorithms, developed for systems whose unknown parameters are physical component values, exploit a special form of partial knowledge, namely that the transfer function polynomials are multilinear in the parameters, to enhance robustness.

Key Words—Adaptive identification; partial knowledge, multilinear; uniform asymptotic convergence.

Abstract—Linear, time-invariant systems with a limited number of unknowns corresponding to physical component values often exhibit certain structural conditions on the way in which these values appear in the transfer functions. In this paper, the problem of identifying these unknowns is considered. The structure translates to an associated identification problem which is multilinear in the unknowns. This is exploited in the design of two algorithms, which use measurements of input and output and knowledge of the polynomial coefficients of the multilinear combinations of the system parameters. Conditions are established for the uniform asymptotic stability of these algorithms.

1. INTRODUCTION

THIS PAPER considers a class of equation error adaptive identification algorithms for systems whose unknown parameters have direct physical relevance. Conventionally, algorithms for identifying linear, time-invariant systems of known finite order (see for example Anderson, 1977a; Carroll and Lindorff, 1973; Kreisselmeier, 1977, 1980; Lion, 1967; Luders and Narendra, 1974; Narendra and Kudva, 1974) presuppose a complete lack of knowledge about the unknown system (aside sometimes from the degree and relative degree) and ignore all additional information available to the modeller. The algorithms thus estimate the numerator and denominator coefficients of the transfer function, having first assumed all of them to be unknown. In practice, however, a great deal of partial knowledge is often available, which if exploited should give rise to parameterizations involving fewer unknowns and better identification schemes.

Usually, the lack of knowledge in a system relates to certain physical parameter values. Thus all parts of a mechanical system may be known *a priori* except perhaps the values of some, but not necessarily all, of its physical elements; e.g. a moment of inertia, a frictional coefficient or the like. Accordingly in the parametrizations considered in this paper, the unknown parameters have direct physical significance.

The number of unknowns in most such cases is small. Moreover, system physics allows assumptions on the parameter magnitude bounds and, in most cases, on the knowledge of their signs. Of the two algorithms formulated in this paper the stability analysis of one, but *not both*, exploits the knowledge of these assumed magnitude bounds.

Dasgupta and Anderson (1987) have shown that when most parts of a linear physical system are known but certain parameters associated with its components are unknown then the system transfer functions of many of these systems are ratios of two polynomials having coefficients multilinear in the unknown parameters. For example, with three unknown parameters, the transfer function has the form

$$T(s, k_1, k_2, k_3) = P(s, k_1, k_2, k_3)/Q(s, k_1, k_2, k_3)$$

* Received 15 July 1985; revised 17 December 1986; received in final form 25 September 1987. The original version of this paper was presented at the 10th IFAC World Congress which was held in Munich, F.R.G. during July 1987. The Published Proceedings of this IFAC Meeting may be ordered from: Pergamon Press plc, Headington Hill Hall, Oxford OX3 0BW, England. This paper was recommended for publication in revised form by Associate Editor T. Soderstrom under the direction of Editor P. C. Parks.

† Department of Electrical and Computer Engineering, University of Iowa, Iowa City, LA 52242, U.S.A.

‡ Department of Systems Engineering, Research School of Physical Sciencies, Australian National University, G.P.O. Box 4, Canberra, A.C.T. 2601, Australia.

§ School of Electrical Engineering and Computer Science, University of New South Wales, P.O. Box 1, Kensington, N.S.W. 23003, Australia.

where

$$P = p_0(s) + k_1 p_1(s) + k_2 p_2(s) + k_3 p_3(s) + k_2 p_{12}(s) + k_2 k_3 p_{23}(s) + k_1 k_3 p_{31}(s) + k_1 k_2 k_3 p_{123}(s)$$

$$Q = q_0(s) + k_1 q_1(s) + k_2 q_2(s) + k_3 q_3(s) + k_1 k_2 q_{12}(s) + k_2 k_2 q_{23}(s) + k_1 k_3 q_{31}(s) + k_1 k_2 k_3 q_{123}(s).$$

Sometimes k_i could be the inverse of the corresponding physical component value and certain summands in P and Q may be zero.

The exceptions are from parameter defining elements like mutual inductors, or gear ratios which allow cross-coupling to occur between different energy storage elements. As an example, consider the dynamics relevant to the attitude control of the communications technology satellite, Hermes (see Diduch and Balasubramanian, 1982). The physical parameters of the associated system are: I_1, the moment of inertia about the roll axis; I_2, that about the yaw axis; ω_0, the orbital rate; h, the nominal wheel angular momentum; β, the offset angle; F_2, the yaw thruster level; L_2, the yaw thruster moment arm; and G_2, the impulse bit factor. The input provides a guide for the level of consumed fuel and the output is the roll. The corresponding transfer function is given by

$$\frac{s\dfrac{F_2 L_2 G_2}{I_2}\left(\dfrac{h}{I_1} - \omega_0\right)}{s^3 + s^2\left(\dfrac{2\omega_0 h}{I_1} + \dfrac{2\omega_0 h}{I_2} - \omega_0^2 - \dfrac{h^2}{I_1 I_2}\right) + \dfrac{\omega_0^2 h^2}{I_1 I_2}}.$$

Notice that coefficients are multilinear in I_1, I_2, F_2, L_2 and G_2, but not in ω_0 and h. But the latter two are the orbital rate and wheel angular momentum, respectively, and thus allow cross-coupling between energy storage devices. Notice also that F_2, L_2 and G_2 are not separately identifiable as they appear only in the product form $F_2 L_2 G_2$. Thus this product must be treated as a single variable. Also for an identification problem it may be necessary to model $1/I_1$ and $1/I_2$ as parameters rather than I_1 and I_2. The advantages of our method are heightened if only a few of the above physical parameters are unknown.

In Section 2, two adaptive identification algorithms which exploit the multilinearity mentioned above are presented. One of these involves a two step procedure. The first step obtains an unconstrained estimate of the parameter vector by ignoring the *a priori* knowledge related to the inherent nonlinearity, while the second step uses this latter knowledge to constrain the parameters on to the desired manifold of the parameter space, i.e. the manifold which takes the multilinearity into account. Simulations presented in this paper show that the second step improves the robustness of the algorithm. The other algorithm essentially combines these two steps into one by using penalty functions. Its guaranteed convergence requires that the bounds on the magnitude of the parameters be known. Given that these parameters are directly related to physical element values, such a requirement is easy to satisfy.

Adaptive algorithms are usually designed and analyzed under certain idealizing assumptions, such as absence of noise and perfect modelling. In all likelihood these assumptions will not hold. The algorithms should be robust enough to withstand modest departures from these assumptions. Usually for such robustness, parameter convergence must occur uniformly asymptotically in the ideal case (see Anderson and Johnstone, 1982; Anderson, 1985). Thus, as uniformly asymptotically stable (u.a.s.) algorithms are totally stable (Hahn, 1967), the algorithms are equipped to overcome moderate deviations from ideality. As with most identification schemes for linear systems, it is shown that uniform asymptotic stability is conditional on (Anderson, 1977a, b; Kreisselmeier, 1977; Morgan and Narendra, 1977; Sondhi and Mitra, 1976) the satisfaction of persistence of excitation (p.e.) conditions. Section 3 derives conditions for uniform asymptotic convergence. In Section 4, simulation results are presented to verify the fact that the second step of the first algorithm improves the overall robustness characteristics.

2. PARAMETER ADJUSTMENT LAWS

2.1. *The system and some notations*

In this paper, we consider the problem of identifying an asymptotically stable SISO system with proper nth order transfer function

$$W(s) = \frac{P(s, \mathbf{k})}{Q(s, \mathbf{k})}. \tag{2.1}$$

Here, $\mathbf{k}$ is an N-vector of the unknown parameters $k_1, \ldots, k_N$ which are to be identified and $P(s, \mathbf{k})$ and $Q(s, \mathbf{k})$ are polynomials in s. Specifically, we study the case where P and Q are multilinear in $k_1, \ldots, k_N$. Thus the transfer function can be written

$$W(s) = \frac{p_0(s) + \sum_{\sigma \in S}\left(\pi_{i \in \sigma} k_i\right) p_\sigma(s)}{q_0(s) + \sum_{\sigma \in S}\left(\pi_{i \in \sigma} k_i\right) q_\sigma(s)}. \tag{2.2}$$

Here, all the $p_{(.)}(s)$ and $q_{(.)}(s)$ are known, and S is defined as follows.

Suppose for any $\sigma \subset \{1, \ldots, N\}$, $p_\sigma(s)$ and $q_\sigma(s)$ are the coefficients of $\pi_{i\in\sigma} k_i$, in the numerator and the denominator, respectively. Then S is the set of all those subsets of $\{1, \ldots, N\}$ such that if $\sigma \in S$ then at least one of $p_\sigma(s)$ and $q_\sigma(s)$ is *not* the zero polynomial. With $|S|$ the number of elements of S, define $\mathbf{K} \in R^{|S|}$ with elements

$$K_\sigma = \pi_{i\in\sigma} k_i, \quad \forall \sigma \in S. \tag{2.3}$$

For example, in (1.1) $\mathbf{K}$ can be $[k_1, k_2, k_3, k_1k_2, k_1k_3, k_2k_3, k_1k_2k_3]^T$. If, however, $p_{31} = q_{31} = 0$ then $\mathbf{K} = [k_1, k_2, k_3, k_1k_2, k_2k_3, k_1k_2k_3]^T$.

In the following, notation will be abused on two scores. The element $K_{\{1,2,3\}}$ for example will be designated by K_{123}. Moreover, $v(s)$ will denote the Laplace transform of $v(t)$ and $b(s)v(t)$ the time function $\mathcal{L}^{-1}[b(s)v(s)]$.

The following assumption will hold throughout the paper.

Assumption 2.1. The vector $\mathbf{K}$ is such that it uniquely defines $\mathbf{k}$.

$\nabla\nabla\nabla$

Thus if $N = 2$, the vector $\mathbf{K} = [k_1, k_1k_2]$ will be acceptable if $k_1 \neq 0$, while $\mathbf{K} = [k_1k_2]$ will not be acceptable.

Denote the input by $u(t)$ and the output by $y(t)$. The input/output description is equivalent to

$$h_0(t) + \mathbf{H}^T(t)\mathbf{K} = 0. \tag{2.4}$$

Here for some $\gamma > 0$

$$h_0(t) = \frac{p_0(s)u(t) - q_0(s)y(t)}{(s+\gamma)^n}$$

and $\mathbf{H}(t)$ has the same dimension as $\mathbf{K}$, having elements

$$h_\sigma(t) = \frac{p_\sigma(s)u(t) - q_\sigma(s)y(t)}{(s+\gamma)^n} \quad \forall \sigma \in S. \tag{2.5}$$

The filter $1/(s+\gamma)^n$ is for avoiding explicit differentiation of u and y. We now present the two step algorithm for estimating $\mathbf{k}$.

2.2. *The two step algorithm*

This algorithm assumes that none of the k_i are zero. This assumption is reasonable because the k_i are physical component values. The two steps in question can be summarized as follows.

Step 1: *ignoring* (temporarily) the nonlinear relationships between the parameters, an unconstrained estimate, referred to as $\mathbf{K}_u$, of $\mathbf{K}$ is produced.

Step 2: using $\mathbf{K}_u$, a constrained estimate $\hat{\mathbf{K}}$ which is in some sense close to $\mathbf{K}_u$ is produced so that it obeys the underlying nonlinear relations.

Both the steps are performed simultaneously.

For example, in a two parameter system with transfer function

$$\frac{p_0(s) + k_1p_1(s) + k_2p_2(s) + k_1k_2p_{12}(s)}{q_0(s) + k_1q_1(s) + k_2q_2(s) + k_1k_2q_{12}(s)} \tag{2.6}$$

the first step generates $\mathbf{K}_u = [K_{u1}, K_{u2}, K_{u12}]^T$ where K_{u12} need not equal $K_{u1}K_{u2}$. The second step generates a $\hat{\mathbf{k}} = [\hat{k}_1, \hat{k}_2]^T$ such that $\hat{\mathbf{K}} = [\hat{k}_1, \hat{k}_2, \hat{k}_1\hat{k}_2]^T$ is in some sense closest to $\mathbf{K}_u$.

In general, $\mathbf{K}_u(t)$ is generated as the state of some differential equation and in the absence of noise or modelling errors,

$$\lim_{t\to\infty} \mathbf{K}_u(t) = \mathbf{K}. \tag{2.7}$$

For example, $\mathbf{K}_u(\cdot)$ might be the solution of

$$\dot{\mathbf{K}}_u(t) = -\beta\mathbf{H}(t)[h_0(t) + \mathbf{H}(t)^T\mathbf{K}_u(t)] \tag{2.8}$$

for some $\beta > 0$.

For the second step the following definitions need to be made. Define $\hat{\mathbf{k}}(t) = [\hat{k}_1(t), \ldots, \hat{k}_N(t)]^T$,

$$\hat{K}_\sigma(t) = \pi_{i\in\sigma} \hat{k}_i(t) \quad \forall \sigma \in S \tag{2.9}$$

and $\hat{\mathbf{K}}(t)$ the vector containing the elements $\hat{K}_\sigma(t)$ with ordering consistent with that of the elements of $\mathbf{K}$. Then the second step is obtained by the steepest descent minimization with respect to $\hat{\mathbf{k}}$ of the quadratic

$$L_u(t) = [\hat{\mathbf{K}}(t) - \mathbf{K}_u(t)]^T \Lambda [\hat{\mathbf{K}}(t) - \mathbf{K}_u(t)] \tag{2.10}$$

where Λ is a positive definite diagonal matrix with diagonal elements $\{\lambda_\sigma\}$, $\sigma \in S$.

Thus with $\hat{\mathbf{k}}(t)$ the estimate of $\mathbf{k}$, $\hat{\mathbf{k}}(t)$, subject to a switching mechanism described below, is obtained from

$$\dot{\hat{\mathbf{k}}}(t) = \frac{-\partial L_u(t)}{\partial \hat{\mathbf{k}}(t)} \tag{2.11}$$

$$= -\left[\frac{\partial \hat{\mathbf{K}}(t)}{\partial \hat{\mathbf{k}}(t)}\right]^T \Lambda (\hat{\mathbf{K}}(t) - \mathbf{K}_u(t)). \tag{2.12}$$

Switching mechanism. There are two rules. (i) If any elements of $\mathbf{K}_u(t)$ have inconsistent signs, e.g. $K_{u12}(t)K_{u1}(t)K_{u2}(t) \leqslant 0$, then (2.12) is temporarily replaced by $\dot{\hat{\mathbf{k}}}(t) = 0$. (ii) If elements of $K_u(t)$ have consistent signs but $\hat{K}_\sigma(t)K_{u\sigma}(t) \leqslant 0$ for some σ, then the $\hat{k}_i(t)$ are reset to any value for which

$$\hat{K}_\sigma(t)K_{u\sigma}(t) > 0 \quad \forall \sigma \in S. \tag{2.13}$$

$\nabla\nabla\nabla$

Given (2.7), it may seem that (2.12) is unnecessary. This is indeed so under conditions of perfect modelling. Suppose, however, there is some noise or modelling error, e.g. nonlinearity or unmodelled high frequency dynamics, present. As such one cannot expect that $\mathbf{K}_u$ will have the correct constrained structure which $\mathbf{K}$ has. Accordingly, our algorithm, by implementing the second step of the two step procedure described above, introduces a gradient descent algorithm to locate the value of $\mathbf{K}$ (satisfying the constraints) which is a best least-squares match of $\mathbf{K}_u$. It can be viewed as an attempt to use the implicit nonlinear relations to filter out the imperfections and to constrain the eventual estimate to lie in the correct model. Simulations presented in Section 4 show how effectively (2.12) ameliorates the inaccuracies in the first step.

2.3. *Single step algorithm*

The second algorithm studied in this paper assumes that $i \in S$, $\forall i \in \{1, \ldots, N\}$, i.e. each k_i is an entry of $\mathbf{K}$. It does not produce an estimate $\hat{\mathbf{k}}(t)$ of $\mathbf{k}$. Instead, it produces an estimate $\mathbf{K}_s(t)$ of $\mathbf{K}$ which tends in the limit to obey the nonlinear relations of (2.3). To achieve this, define the function

$$L_s(\mathbf{k}_s, \mathbf{k}) = \tfrac{1}{2}\Big[\lambda(\mathbf{K}_s - \mathbf{K})^{\mathrm{T}}(\mathbf{K}_s - \mathbf{K}) + \sum_{\sigma \in S} \lambda_\sigma \Big(K_{s\sigma} - \prod_{i \in \sigma} K_{si}\Big)^2\Big] \quad (2.14)$$

where $\lambda > 0$ and $K_{s\sigma}$ is the σ-th entry of $\mathbf{K}_s$ and K_{si} is the $\{i\}$-th entry. Thus the first term in (2.14) is the square of the Euclidean distance between $\mathbf{K}_s$ and $\mathbf{K}$ and the second term is a measure of the amount by which $\mathbf{K}_s$ fails to obey the nonlinear relations implied by (2.3).

Observe that with

$$\mathbf{r}(t) \triangleq \int_0^t e^{-\alpha(t-\tau)} h_0(\tau)\mathbf{H}(\tau)\, d\tau \quad (2.15)$$

and

$$\mathbf{R}(t) \triangleq \int_0^t e^{-\alpha(t-\tau)} \mathbf{H}(\tau)\mathbf{H}(\tau)^{\mathrm{T}}\, d\tau \quad (2.16)$$

for some $\alpha > 0$, (2.4) can be re-expressed as

$$\mathbf{R}(t)\mathbf{K} + \mathbf{r}(t) = 0. \quad (2.17)$$

Then the single step algorithm is

$$\dot{\mathbf{K}}_s(t) = -\mathbf{R}(t)\frac{\partial L_s}{\partial \mathbf{K}_s}(\mathbf{K}_s(t), \mathbf{k}). \quad (2.18)$$

We shall now show how (2.18) can be implemented.

Define the vector $\chi_\sigma(\mathbf{K}_s)$ in $R^{|S|}$ to be

$$\chi_\sigma(\mathbf{K}_s) \triangleq \Big[\frac{\partial}{\partial \mathbf{K}_s}\Big(K_{s\sigma} - \prod_{i \in \sigma} K_{si}\Big)\Big]. \quad (2.19)$$

For example, for $N = 2$ and $\sigma = \{1, 2\}$

$$\chi_\sigma(\mathbf{K}_s) = [-K_{s2}, -K_{s1}, 1]^{\mathrm{T}}.$$

Then, (2.18) can be implemented as

$$\dot{\mathbf{K}}_s(t) = -\lambda[\mathbf{R}(t)\mathbf{K}_s(t) + \mathbf{r}(t)] - \mathbf{R}(t)\sum_{\sigma \in S}\Big\{\lambda_\sigma\Big(K_{s\sigma} - \prod_{i \in \sigma} K_{si}\Big)\chi_\sigma(\mathbf{K}_s(t))\Big\}. \quad (2.20)$$

This follows by differentiation of (2.14) and observing that

$$\mathbf{r}(t) = -\mathbf{R}(t)\mathbf{K}. \quad (2.21)$$

3. CONVERGENCE ANALYSIS

In this section the parameter adjustment algorithms presented in Section 2 are analyzed for convergence. The standing assumptions for this section are Assumption 2.1 and the fact that none of the k_i are zero. The justification for the latter assumption has been given in Section 2. At the outset the following lemma is needed.

Lemma 3.1. With Λ, $\hat{\mathbf{K}}$, $\hat{\mathbf{k}}$, $\mathbf{k}$ and $\mathbf{K}$ defined in Section 2 consider

$$L = \tfrac{1}{2}(\hat{\mathbf{K}} - \mathbf{K})^{\mathrm{T}}\Lambda(\hat{\mathbf{K}} - \mathbf{K}). \quad (3.1)$$

Suppose the standing assumptions hold and $\forall \sigma \in S$

$$\hat{K}_\sigma K_\sigma > 0. \quad (3.2)$$

Then

$$\frac{\partial L}{\partial \hat{\mathbf{k}}} = 0 \quad \text{iff } \hat{\mathbf{k}} = \mathbf{k}. \quad (3.3)$$

Proof. The if part is trivial. For the only if part observe that the ith element of $\partial L/\partial \hat{\mathbf{k}}$ is

$$\Big(\frac{\partial L}{\partial \hat{\mathbf{k}}}\Big)_i = \sum_{\sigma \in S} \lambda_\sigma \Big\{\prod_{\substack{j \in S \\ j \neq i}} \hat{k}_j\Big\}\Big[\Big(\prod_{j \in \sigma} \hat{k}_j\Big) - \Big(\prod_{j \in \sigma} k_j\Big)\Big]. \quad (3.4)$$

Suppose $\partial L/\partial \hat{\mathbf{k}} = 0$. Then for arbitrary N-dimensional η

$$\eta^{\mathrm{T}}\frac{\partial L}{\partial \hat{\mathbf{k}}} = 0. \quad (3.5)$$

Choose

$$\eta^{\mathrm{T}} = \phi^{\mathrm{T}}\,\mathrm{diag}\,\{\hat{k}_i\}_{i=1}^N \quad (3.6)$$

with

$$\phi_i = \ln\frac{\hat{k}_i}{k_i}. \quad (3.7)$$

Note (3.2) and the standing assumptions ensure that ϕ_i is defined. Then

$$\begin{aligned}\eta^T\frac{\partial L}{\partial \hat{\mathbf{k}}} &= \sum_{\sigma\in S}\sum_{i\in\sigma}\lambda_\sigma\left\{\ln\frac{\hat{k}_i}{k_i}\right\}\\ &\quad\times\hat{k}_i\Big(\prod_{\substack{j\in\sigma\\ j\neq i}}\hat{k}_j\Big)\Big\{\prod_{j\in\sigma}\hat{k}_j-\prod_{j\in\sigma}k_j\Big\}\\ &= \sum_{\sigma\in S}\lambda_\sigma\Big(\prod_{j\in\sigma}\hat{k}_j\Big)\Big\{\prod_{j\in\sigma}\hat{k}_j-\prod_{j\in\sigma}k_j\Big\}\\ &\quad\times\sum_{i\in\sigma}\ln\frac{\hat{k}_i}{k_i}\\ &= \sum_{\sigma\in S}\lambda_\sigma\Big\{\Big(\prod_{j\in\sigma}\hat{k}_j\Big)^2-\Big(\prod_{j\in\sigma}k_j\Big)\Big(\prod_{j\in\sigma}\hat{k}_j\Big)\Big\}\\ &\quad\times\ln\frac{\Big(\prod_{j\in\sigma}\hat{k}_j\Big)^2}{\Big(\prod_{j\in\sigma}k_j\Big)\Big(\prod_{j\in\sigma}\hat{k}_j\Big)}\\ &= \sum_{\sigma\in S}\lambda_\sigma\alpha_\sigma\beta_\sigma \qquad (3.8)\end{aligned}$$

with $\alpha_\sigma=\Big(\prod_{j\in\sigma}\hat{k}_j\Big)^2-\Big(\prod_{j\in\sigma}\hat{k}_j\Big)\Big(\prod_{j\in\sigma}k_j\Big)$ and β_σ obviously defined. Now suppose $\alpha_\sigma\neq 0$. Then

$$\begin{aligned}&\alpha_\sigma\gtrless 0\\ &\Leftrightarrow\Big(\prod_{j\in\sigma}\hat{k}_j\Big)^2\lesseqgtr\Big(\prod_{j\in\sigma}\hat{k}_j\Big)\Big(\prod_{j\in\sigma}k_j\Big)\\ &\Leftrightarrow\Big(\prod_{j\in\sigma}\hat{k}_j\Big)^2\Big/\Big(\prod_{j\in\sigma}\hat{k}_j\Big)\Big(\prod_{j\in\sigma}k_j\Big)\lesseqgtr 1\\ &\Leftrightarrow\beta_\sigma\lesseqgtr 0. \qquad (3.9)\end{aligned}$$

Hence (3.5) implies $\alpha_\sigma=0\ \forall\sigma\in S$, whence by the definition of α_σ and Assumption 2.1, the result follows.

∇∇∇

This result will be used in different contexts to prove the convergence of the two algorithms.

3.1. *Convergence of the two step algorithm*

For the two step algorithm we shall assume that the first step is any adaptive identifier satisfying the following assumption.

Assumption 3.1. The unconstrained estimate is generated in such a way that there exists a Lyapunov function L_1 in the error $\mathbf{K}_u(t)-\mathbf{K}$ such that

$$c_1\|\mathbf{K}_u(t)-\mathbf{K}\|^2\leqslant L_1\leqslant c_2\|\mathbf{K}_u(t)-\mathbf{K}\|^2$$

and

$$\dot{L}_1\leqslant -c_3\|\mathbf{K}_u(t)-\mathbf{K}\|^2$$

where c_1, c_2 and $c_3>0$.

∇∇∇

Assumption 3.1 implies the exponential asymptotic convergence of $\mathbf{K}_u(t)-\mathbf{K}$ to zero (Krasovskii, 1963, p. 86).

For example, if (2.8) is used then the assumption is satisfied provided for some α_1, α_2 and $\delta>0$ and all σ

$$\alpha_1\mathbf{I}\leqslant\int_\sigma^{\sigma+\delta}\mathbf{H}(t)\mathbf{H}^T(t)\,dt\leqslant\alpha_2\mathbf{I}. \qquad (3.10)$$

To prove convergence of the overall algorithm, we need the following lemma.

Lemma 3.2. Suppose Assumption 3.1 and the standing assumptions hold. Then $\hat{\mathbf{k}}(t)$ is bounded and there exists T independent of the initial time t_0, such that $\forall t\geqslant t_0+T$, and σ

$$K_\sigma\hat{K}_\sigma(t)>0. \qquad (3.11)$$

Further, no switchings in $\hat{\mathbf{k}}(t)$ take place beyond $t=T$.

Proof of Lemma 3.2. By Assumption (3.1), $\mathbf{K}_u-\mathbf{K}$ is exponentially convergent. Thus there exists a T, independent of t_0, for which $\forall t\geqslant t_0+T$ and some $\varepsilon>0$

$$K_{u\sigma}(t)K_\sigma>\varepsilon. \qquad (3.12)$$

Thus the switching mechanism outlined before implies that no switchings according to the first rule take place beyond $t=t_0+T$. Also, $\forall i\in\{1,\ldots,N\}$, (3.12) and switching rule (ii) imply

$$k_i\hat{k}_i(t_0+T)>0. \qquad (3.13)$$

Note that the exponential convergence of $\mathbf{K}_u(t)-\mathbf{K}$ implies that $\mathbf{K}_u(t)$ is bounded. Using (2.12) and the boundedness of $\mathbf{K}_u(t)$, some manipulations will show that $\hat{k}_i\dot{\hat{k}}_i$ is necessarily negative if $|\hat{k}_i|$ is large enough. This means that $\hat{k}_i(t)$ is bounded.

Suppose (3.11) is violated. Then in view of (3.13) and the continuity of the trajectory of $\hat{\mathbf{k}}(t)$, for some t_1, at least one of $|\hat{k}_i(t_i)|$ must become arbitrarily small while $\hat{k}_j(t_1)k_j(t_1)>0$ $\forall j\in\{1,\ldots,N\}$. Then by (2.12)

$$\begin{aligned}k_i\dot{\hat{k}}_i(t_1) &= -\sum_{\substack{\sigma\in S\\ i\in\sigma}}\lambda_\sigma k_i\Big(\prod_{\substack{j\in\sigma\\ j\neq i}}\hat{k}_j(t_1)\Big)\\ &\quad+\Big(\prod_{j\in\sigma}\hat{k}_j(t_1)-K_{u\sigma}(t_1)\Big)\\ &= -\sum_{\substack{\sigma\in S\\ i\in\sigma}}\lambda_\sigma k_i\prod_{\substack{j\in\sigma\\ j\neq i}}\hat{k}_j(t_1)\prod_{j\in\sigma}\hat{k}_j(t)\\ &\quad+\sum_{\substack{\sigma\in S\\ i\in\sigma}}\lambda_\sigma k_i\prod_{\substack{j\in\sigma\\ j\neq i}}\hat{k}_j(t_1)K_{u\sigma}(t_1). \qquad (3.14)\end{aligned}$$

Since $\hat{k}_i(t_1)$ is arbitrarily small, the first summand is less in magnitude than the second. The latter has positive sign by (3.13) and the fact that

$\hat{k}_j(t_1)k_j(t_1) > 0$. Hence the right side of (3.14) is positive, $k_i\hat{k}_i$ will increase and (3.11) will not be violated.

∇∇∇

Theorem 3.1. Suppose Assumption 3.1 and the standing assumption hold. Then $[(\mathbf{K}_u(t) - \mathbf{K})^T, (\hat{\mathbf{k}}(t) - \mathbf{k})^T]^T$ is u.a.s.

Proof. By Lemma (3.2), $\hat{\mathbf{k}}(t)$ is bounded and for some T independent of t_0, (3.11) holds. Thus in view of Lemma 3.1, $\forall t \geq t_0 + T$, $\|\partial L/\partial \hat{\mathbf{k}}(t)\|_2$ is positive definite in $\hat{\mathbf{k}}(t) - \mathbf{k}$. Note that due to (2.12), (3.1) and Lemma 3.2

$$\dot{\hat{\mathbf{k}}}(t) = -\left[\frac{\partial \hat{\mathbf{K}}}{\partial \hat{\mathbf{k}}}\right]^T \Lambda(\hat{\mathbf{K}} - \mathbf{K}) + \left[\frac{\partial \hat{\mathbf{K}}(t)}{\partial \hat{\mathbf{k}}(t)}\right]^T \Lambda(\mathbf{K}_u(t) - \mathbf{K})$$
$$= -\frac{\partial L}{\partial \hat{\mathbf{k}}(t)} + \left[\frac{\partial \hat{\mathbf{K}}(t)}{\partial \hat{\mathbf{k}}(t)}\right]^T \Lambda(\mathbf{K}_u(t) - \mathbf{K}). \quad (3.14a)$$

Due to the boundedness of $\hat{\mathbf{k}}(t)$, there exists a finite M such that

$$\left\|\frac{\partial \hat{\mathbf{K}}^T}{\partial \hat{\mathbf{k}}}\Lambda\right\| \leq M.$$

Choose the Lyapunov function $L + \beta L_1$ with L_1 defined in Assumption 3.1 and a positive β to be specified presently. Then

$$\dot{L} + \beta\dot{L}_1 = -\left\|\frac{\partial L}{\partial \hat{\mathbf{k}}}\right\|^2 + \left[\frac{\partial L}{\partial \hat{\mathbf{k}}}\right]^T \times \left[\frac{\partial \hat{\mathbf{K}}(t)^T}{\partial \hat{\mathbf{k}}(t)}\Lambda(\mathbf{K}_u(t) - \mathbf{K})\right] + \beta\dot{L}_1$$
$$\leq -\left\|\frac{\partial L}{\partial \hat{\mathbf{k}}}\right\|^2 + M\left\|\frac{\partial L}{\partial \hat{\mathbf{k}}}\right\| \times \|\mathbf{K}_u(t) - \mathbf{K}\| - \beta c_3 \|\mathbf{K}_u(t) - \mathbf{K}\|^2$$
$$= -\left(\left\|\frac{\partial L}{\partial \hat{\mathbf{k}}}\right\| - \frac{M}{2}\|\mathbf{K}_u(t) - \mathbf{K}\|\right)^2 + \left(\frac{M^2}{4} - \beta c_3\right)\|\mathbf{K}_u(t) - \mathbf{K}\|^2.$$

Then for $\beta > M^2/4c_3$, $\dot{L} + \beta\dot{L}_1$ is negative definite. Thus the result follows.

3.2. *Convergence of the single step law*

Consider the differential equation (2.18),

$$\dot{\mathbf{K}}_s = -\mathbf{R}(t)\left[\frac{\partial L_s}{\partial \mathbf{K}_s}(\mathbf{K}_s(t), \mathbf{k})\right]$$

with L_s defined in (2.14). Assume that $K_i = k_i$ $\forall i \in \{1, \ldots, N\}$.

The following theorem describes conditions for its uniform asymptotic stability.

Theorem 3.2. The adaptive law in (2.18) is u.a.s. in the large if $\forall t \geq t_0$

(i) $K_{si}(t)k_i > 0 \quad \forall i \in \{1, \ldots, N\}$ (3.15)

and

(ii) there exists $\alpha_1 > 0$ such that for some $\bar{t}$ and all $t > \bar{t}$

$$\mathbf{R}(t) > \alpha_1 \mathbf{I}. \quad (3.16)$$

Proof. Consider the candidate Lyapunov function L_s defined in (2.14). Now, L_s is positive definite with $L_s = 0$ iff $\mathbf{K}_s = \mathbf{K}$. From (2.18)

$$-\dot{L}_s(t) = \frac{\partial L_s^T}{\partial \mathbf{K}_s}\mathbf{R}(t)\frac{\partial L_s}{\partial \mathbf{K}_s} \geq \alpha_1\left\|\frac{\partial L_s}{\partial \mathbf{K}_s}\right\|_2^2. \quad (3.17)$$

Since $[\partial L_s/\partial \mathbf{K}_s]$ has no explicit dependence on t, uniform asymptotic stability follows if

$$\frac{\partial L_s}{\partial \mathbf{K}_s} = 0 \Rightarrow \mathbf{K}_s = \mathbf{K}. \quad (3.18)$$

From $\partial L_s/\partial K_{s\sigma} = 0$ for $|\sigma| > 1$ it follows that

$$K_{s\sigma} = \left(\lambda\prod_{i\in\sigma} k_i + \lambda_\sigma \prod_{i\in\sigma} K_{si}\right)\Big/(\lambda + \lambda_\sigma). \quad (3.19)$$

Setting $\partial L_s/\partial K_{si} = 0$ for all $i \in S$, and using (3.19), yields after some manipulation

$$\lambda(K_{si} - k_i) + \sum_{\substack{\sigma\in S\\ i\in\sigma}} \frac{\lambda_\sigma\lambda}{\lambda_\sigma + \lambda}\left(\prod_{\substack{j\in\sigma\\ i\in\sigma}} K_{sj}\right) \times \left(\prod_{j\in\sigma} K_{sj} - \prod_{j\in\sigma} k_j\right) = 0, \quad \forall i \in \{1, \ldots, N\}. \quad (3.20)$$

Now in (3.1) substitute K_{si}, $i \in \{1, \ldots, N\}$, for $\hat{k}_i$. Then (3.20) is precisely

$$\left.\frac{\partial L}{\partial \hat{\mathbf{k}}}\right|_{\hat{\mathbf{k}}_i = \bar{\mathbf{K}}_{si}}$$

where $\bar{\mathbf{K}}_s$ is in R^N and has the first N elements of $\mathbf{K}_s$. Of course the Λ matrix in the definition of L in (3.1) should be modified appropriately, viz.

$$\lambda_i = \lambda \quad \forall i \in \{1, \ldots, N\}$$

and

$$\lambda_\sigma = \frac{\lambda_\sigma\lambda}{\lambda_\sigma + \lambda} \quad \forall \sigma \in S \text{ and } \sigma \notin \{1, \ldots, N\}.$$

Thus in view of (3.15) and Lemma 3.1, (3.20) implies that

$$K_{si} = k_i \quad \text{for all } i \in \{1, \ldots, N\}. \quad (3.21)$$

Substitution into (3.19) then yields

$$K_{s\sigma} = \prod_{i\in\sigma} k_i \quad \text{for all } \sigma \in S \quad (3.22)$$

and thus (3.18) holds.

∇∇∇

To ensure that (3.15) holds we use the known magnitude bounds on the k_i and a linear translation in the parameters. Let us illustrate with a two parameter example.

Suppose k_i are known *a priori* to lie in a region A, defined by

$$m_i \leq k_i \leq M_i \quad i \in \{1, 2\}.$$

Define $\bar{k}_i = k_i + \Delta$ for a suitable $\Delta > 0$. Call the region containing the $\bar{k}_i$, region B. Choose Δ such that $m_i + \Delta > 0$.

By (3.17)

$$L_s(t) \leq L_s(0) \quad \forall t \in R_+. \tag{3.22a}$$

Select $\mathbf{K}_s(0)$ so that $[K_{S1}(0), K_{S2}(0)]^T \in B$ and $K_{s_{12}}(0) = K_{s_1}(0)K_{s_2}(0)$ with $\mathbf{K}_s$ now identifying $\bar{\mathbf{K}} \triangleq [\bar{k}_1, \bar{k}_2, \bar{k}_1\bar{k}_2]^T$. Then with $\Delta_i = M_i - m_i$ we have

$$L_s(0) \leq \frac{\lambda}{2}(\Delta_1^2 + \Delta_2^2) + \frac{\lambda}{2}\{(\Delta_1 + \Delta_2)(m_1 + \Delta) + |m_2 - m_1|\Delta_1 + \Delta_1\Delta_2\}^2. \tag{3.23}$$

Suppose for some t_1, $K_{s_1}(t_1) \geq 0$. Then, with reasoning set out below,

$$L_s(t_1) \geq \frac{\lambda}{2}\bar{k}_1^2 + \frac{\lambda}{2}(K_{s_{12}} - \bar{k}_1\bar{k}_2)^2 + \frac{\lambda_{12}}{2}K_{s_{12}}^2 \tag{3.23a}$$

$$\geq \frac{\lambda}{2}\bar{k}_1^2 + \frac{1}{2}\left(\frac{\lambda\lambda_{12}^2}{(\lambda + \lambda_{12})^2} + \frac{\lambda_{12}\lambda^2}{(\lambda + \lambda_{12})^2}\right)\bar{k}_1^2\bar{k}_2^2 \tag{3.23b}$$

$$= \frac{\lambda}{2}\bar{k}_1^2 + \frac{\lambda\lambda_{12}}{2(\lambda + \lambda_{12})}\bar{k}_1^2\bar{k}_2^2$$

$$\geq \frac{\lambda}{2}(m_1 + \Delta)^2 + \frac{\lambda\lambda_{12}}{2(\lambda + \lambda_{12})} \times (m_1 + \Delta)^2(m_2 + \Delta)^2. \tag{3.24}$$

Here, (3.23a) follows from (2.14) and positivity of $(\lambda/2)(\bar{k}_2 - K_{s_2})^2$, (3.23b) follows by minimizing the right side of (3.23a) with respect of $K_{s_{12}}$ and (3.24) by the fact that $\bar{k}_i \geq m_i + \Delta$. The highest powers of Δ in (3.23) and (3.24) are 2 and 4, respectively. Thus we can find a large enough Δ so that the right side of (3.23) is smaller than (3.24). For such a choice of Δ, when $K_{s_1}(t_1) = 0$,

$$L_s(t_1) > L_s(0)$$

which contradicts (3.22a). A similar argument shows that for large enough Δ, $K_{s_2}(t_1) = 0$ implies

$$L_s(t) > L_s(0).$$

Thus in view of (3.22a), a translation of the parameters by a suitable amount, along with an appropriate initialization will ensure that (3.15) holds. Accordingly in such a case the single step algorithm is u.a.s. whenever (3.16) holds.

3.3. *Persistence of excitation*

Central to the convergence conditions of any adaptive algorithm is a persistence of excitation (p.e.) requirement. Such a requirement is present here as well. For the two step algorithm it comes in the form of a sufficient condition for the satisfaction of Assumption 3.1. For example, if (2.8) is the first step then (3.10) is necessary and sufficient (see Anderson, 1977) for this assumption to hold. Thus (3.10) will be the p.e. condition for such a choice of the first step. For the single step algorithm the p.e. condition is (3.16).

Conditions such as (3.10) and (3.16) involve $u(t)$ and $y(t)$. For design purposes one would like to have a condition on $u(t)$ only. The problem of such a design has been addressed in Dasgupta *et al.* (1988).

4. SIMULATION RESULTS

In this section we present simulation results for the two step adaptive identifier. These show that the inclusion of the second step does indeed improve performance in the face of a whole range of deviations from ideality. The system considered has transfer function

$$T_p(s) = \frac{(s+2)(s+3)}{(s+1)(s+4)(s+5)} \tag{4.1}$$

parameterized as

$$T_p(s) = \frac{s^2 + k_2 s + k_2}{(s^3 + 6s^2 - s - 10) + 2k_1 s^2 - k_2 s^2 + k_1 k_2 (s+1)} \tag{4.2}$$

with $k_1 = 5$ and $k_2 = 6$.

By results in Dasgupta *et al.* (1988) it can be seen that p.e. is guaranteed whenever the input has two distinct frequencies. This requirement is satisfied in all the simulations. All the λ_σs are unity, unless otherwise specified.

The unconstrained estimate $\mathbf{K}_u(t)$ is generated using a recursive least-squares procedure, with an exponential forgetting factor. The output is generated by simulating (4.2) with finite initial conditions.

4.1. *Slow sinusoidal drift in the parameters*

To assess the ability of the second step to track slow changes in the plant, the parameter vector $\mathbf{k}$ is varied sinusoidally in time as $k_1 = 5 + 0.5 \sin 0.01t$ and $k_2 = 6 + 0.5 \sin 0.01t$. We need to emphasize at this point that the

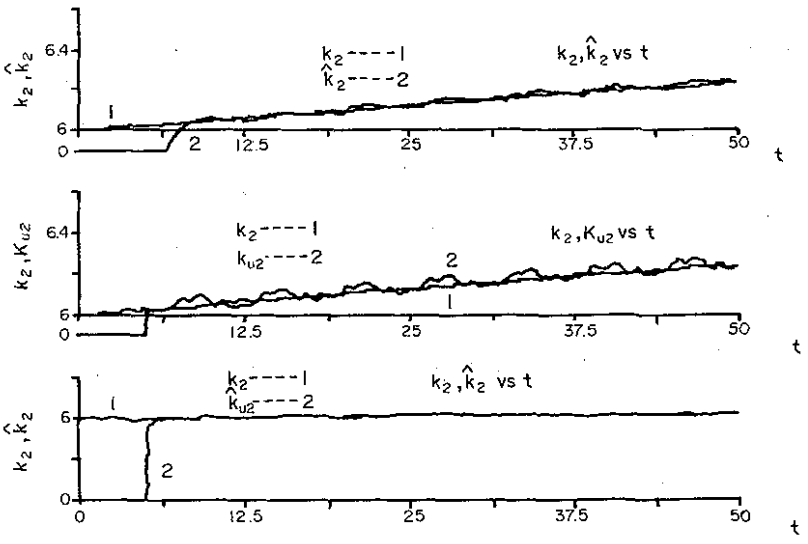

FIG. 1. Sinusoidal drift in the k_i.

choice of a sinusoidal function is not critical. What is important is that the parameters drift slowly and remain bounded. Here "slowness" is with respect to the time constants of the system and the exponential forgetting factor. The improvement gained from the second step can be seen from Fig. 1, which shows that $\hat{k}_2$ tracks k_2 more accurately than does K_{u2}.

4.2. *Identification when system has high frequency unmodelled modes*

The behavior of the identification algorithm for the case where the plant is not a member of the model set is examined by introducing an unmodelled mode at $-100 \pm 14j$. Thus it has a transfer function

$$T_{pu}(s) = \frac{10196(s+2)(s+3)}{(s+1)(s+4)(s+5)(s^2+200s+10196)}$$

although the identifier is designed for (4.2). Figure 2 shows that $\hat{k}_2$ tracks the value 5 better than K_{u2} does and also that the trajectory is smoother. The choice of input was as before, though the frequencies were chosen to be much smaller than the unmodelled poles.

4.3. *Biased noise in the unconstrained parameter estimate*

In this subsection, we assume that $\mathbf{K}_u$ is directly available as $[5, 6, 30]^T + \boldsymbol{\eta}(t)$ where, for $i = 1, 2, 3$, $E[\eta_i(t)] = 0.5$ and $\mathrm{Var}[\eta_i(t)] = 1$. This $\mathbf{K}_u(t)$ is now fed into the second step. The effect of the second step can be seen (see Fig. 3) to be both smoothing and give a significant reduction in bias. A number of simulations with different values of λ_{12}, including zero, indicate that the smoothing, which varied little with λ_{12}, resulted from the low pass filter behavior of the second step.

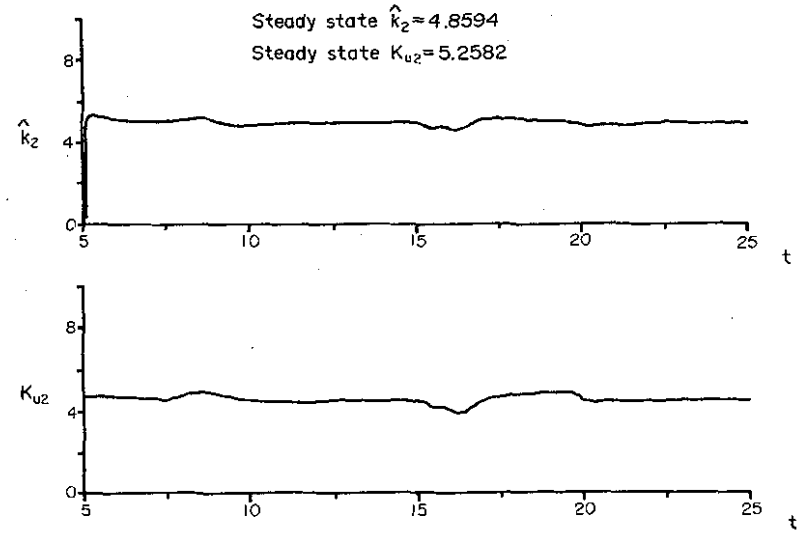

FIG. 2. Identification with unmodelled modes.

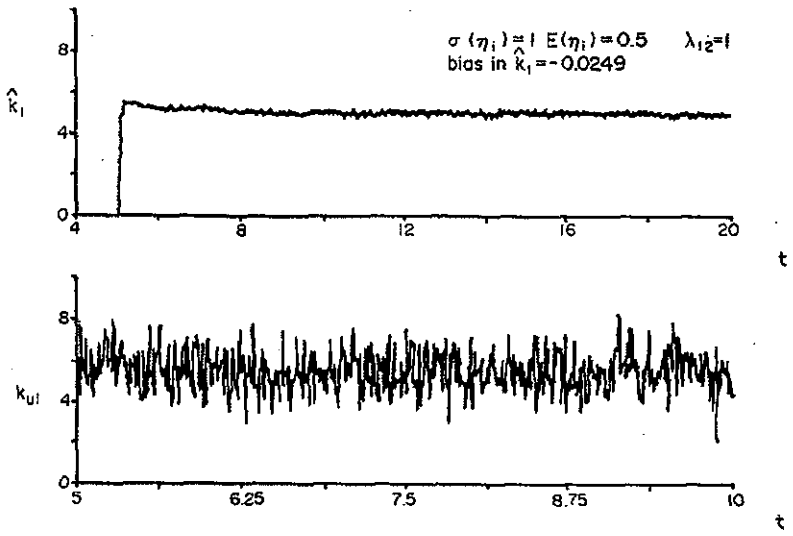

FIG. 3. Biased noise in K_u.

5. CONCLUSIONS

In this paper, the problem of identifying linear time-invariant systems with a limited number of unknown physical parameter values affecting the system transfer function in a multilinear fashion, is studied.

Two equation error identification schemes are presented. One of these is a two step procedure. The first step produces an estimate of the polynomial coefficients by ignoring the multilinear dependence amongst them while the second step uses this dependence to estimate the parameter values. The second scheme performs both functions in a single step by using a penalty function approach.

Both algorithms are shown to be uniformly asymptotically stable under certain persistence of excitation conditions.

Numerical tests of the two step procedure demonstrate the ability of the second step to reduce the effects of biased noise in the unconstrained coefficient estimate. Further, the algorithm exhibits a good ability to track slow variations in parameters and performs well in the presence of high frequency unmodelled modes. This good quality performance in non-ideal circumstances was an expected result of the uniform asymptotic nature of the convergence in the ideal case.

REFERENCES

Anderson, B. D. O. (1977a). An approach to multivariable system identification. *Automatica*, **13**, 401–408.

Anderson, B. D. O. (1977b). Exponential stability of linear equations arising in adaptive identification. *IEEE Trans. Aut. Control*, **AC-22**, 83–88.

Anderson, B. D. O. (1985). Adaptive systems, lack of persistence of excitation and bursting phenomena. *Automatica*, **21**, 247–259.

Anderson, B. D. O. and R. M. Johnstone (1982). When will adaptive systems really adapt? The robustness issue. *Proc. 2nd Conf. on Control Engng*, I.E. Aust., pp. 59–66.

Carroll, R. L. and D. P. Lindorff (1973). An adaptive observer for single input single output linear systems. *IEEE Trans. Aut. control*, **AC-18**, 428–435.

Dasgupta, S. and B. D. O. Anderson (1987). Physically based parameterizations for designing adaptive algorithms. *Automatica*, **23**, 469–477.

Dasgupta, S., B. D. O. Anderson and R. J. Kaye (1988). Persistence of excitation conditions for partially known system. *Automatica*, to appear.

Diduch, C. and R. Balasubramanian (1982). Integrated attitude control of a satellite using a microprocessor. *Proc. IFAC Symp. on Theory and Application of Digital control*, pp. 9–14, New Delhi, January 1982.

Hahn, W. (1967). *Stability of Motion*. Springer, Berlin.

Krasovskii, N. N. (1963). *Stability of Motion*. Stanford University Press.

Kreisselmeier, G. (1977). Adaptive observers with exponetial rate of convergence. *IEEE Trans. Aut. Control*, **AC-22**, 2–8.

Kreisselmeier, G. (1980). Algebraic separation in realizing a linear state feedback control law by means of adaptive observers. *IEEE Trans. Aut. Control*, **AC-25**, 238–243.

Lion, P. M. (1967). Rapid identification of linear and nonlinear systems. *AIAA J*, **5**, 1835–1842.

Luders, G. and K. S. Narendra (1974) Stable adaptive schemes for estimation and identification of linear systems. *IEEE Trans. Aut. Control*, **AC-19**, 841–847.

Morgan, A. P. and K. S. Narendra (1977). On the uniform asymptotic stability of certain linear non-autonomous differential equations. *SIAM J. Control*, **15**, 5–24.

Narendra, K. S. and P. Kudva (1974). Stable adaptive schemes for system identification and control—Parts I and II. *IEEE Trans. Syst. Man. Cybern.*, **SMC-4**, 542–560.

Sondhi, M. M. and D. Mitra (1976). New results on the performance of a well known class of adaptive filters. *Proc. IEEE*, **64**, 1583–1597.