ADAPTIVE CONTROL OF SYSTEMS WITH UNKNOWN PHYSICAL ELEMENT VALUES

SOURA DASGUPTA
Department of Electrical and Computer Engineering, University of Iowa, Iowa City, IA 52242, U.S.A.

BRIAN D. O. ANDERSON
Department of Systems Engineering, Australian National University, GPO Box 4, Canberra, ACT 2601, Australia

AND

R. J. KAYE
School of Electrical Engineering and Computer Science, University of New South Wales, Kensington, NSW 2033, Australia.

SUMMARY
An adaptive control algorithm is proposed for systems with unknown physical element values. These parameters enter the numerator and denominator polynomials of the system transfer function in a multilinear fashion. The algorithm developed is shown to be globally stable with uniformly asymptotic parameter convergence.

KEY WORDS Adaptive control Multilinear Physical parameters Convergence

1. INTRODUCTION
This paper considers the indirect adaptive control of systems with physical element values as the unknown parameters. Conventionally, both direct\(^1\)–\(^3\) and indirect\(^4\),\(^5\) controllers for linear, time-invariant, finite-order systems presuppose a complete lack of knowledge about the unknown system, aside sometimes from the order and relative order. Consequently, they implicitly estimate the numerator and denominator coefficients of the system transfer function, having assumed them all to be unknown.

Often, the only unknowns in a system are a small number of physical parameter values. Thus all parts of a mechanical system may be known \textit{a priori,} except perhaps the value of a moment of inertia, a frictional coefficient or the like. Accordingly, in the underlying parametrization of this paper the unknown parameters have direct physical significance. In any given system, usually such parameters are few in number, certainly fewer than the transfer function coefficients. Thus adaptive algorithms involving such parametrizations can be expected to comprise fewer update equations.

When most parts of a linear, time-invariant, finite-dimensional system are known but certain physical parameters are unknown, Dasgupta and Anderson\(^6\) have shown the numerator and
denominator coefficients of the associated transfer function to be multilinear in these parameter values. Thus, if only two physical parameters with values \( k_1 \) and \( k_2 \) are unknown, the transfer function will be

\[
\frac{p_0(s) + k_1 p_1(s) + k_2 p_2(s) + k_1 k_2 p_{12}(s)}{q_0(s) + k_1 q_1(s) + k_2 q_2(s) + k_1 k_2 q_{12}(s)}
\]

(1)

where the \( p_\sigma(s) \) and \( q_\sigma(s) \) are known polynomials in \( s \). The exceptions arise from elements such as gyrators, gear ratios and mutual inductors, since they allow cross-coupling of energy storage devices.

This paper considers an algorithm for controlling systems typified by (1). The algorithm is similar to that of Kreisselmeier.\(^4\) It thus involves the coupling of a general controller, satisfying certain weak assumptions, with a specific identification algorithm. The identifier conforms to a two-step strategy postulated by Dasgupta \textit{et al.}\(^7-9\). The first step obtains an unconstrained estimate of the modelled parameters by ignoring the inherent multilinearity. The second step constrains the parameters to the manifold associated with this multilinearity and, as simulations show, improves the accuracy of the estimated parameters and performance in the presence of modelling inadequacies. The algorithm and its convergence are discussed in Section 2.

The uniform asymptotic convergence of the parameter estimates and the boundedness of all the signals appearing in the closed loop are conditioned upon a persistence-of-excitation condition on the reference input. This condition is derived in Section 3, which also describes certain implications for robustness associated with this issue.

2. A RECURSIVE-LEAST-SQUARES-BASED ADAPTIVE CONTROLLER

Consider a plant described by the \( n \)th-order, strictly proper transfer function

\[
T_p(s) = \frac{p_0(s) + \sum_{\sigma \subseteq S} \left( \prod_{i \in \sigma} k_i \right) p_\sigma(s)}{q_0(s) + \sum_{\sigma \subseteq S} \left( \prod_{i \in \sigma} k_i \right) q_\sigma(s)}
\]

(2)

where \( k \in \mathbb{R}^N \) is the unknown parameter vector and the \( p_\sigma \) and \( q_\sigma \) are known polynomials. Usually \( N \) is small.

Throughout, notation will be abused by referring, for example, to \( y(s) \) as the Laplace transform of \( y(t) \). Quantities such as \( y_\gamma(t) = (a(s)/b(s))u(t) \) will refer to the solutions of the differential equation \( b(p)y_\gamma(t) = a(p)u(t) \) with \( p = d/dt \) and arbitrary, finite initial conditions. Also, for \( \sigma = [1,2] \), for example, \( p_\sigma(s) \) will be referred to as \( p_{12}(s) \). Let \( K \) be a vector containing the multilinear combinations of the \( k_i \) and let \( K_u \) and \( \hat{K} \) be respectively the unconstrained and constrained estimates of \( K \) in the sense described in Section 1. Thus for \( N = 2 \), \( K = [k_1, k_2, k_1 k_2]^T \), \( K_u = [K_{u11}, K_{u12}, K_{u12}] \) and \( \hat{K} = [\hat{k}_1, \hat{k}_2, \hat{k}_1 \hat{k}_2]^T \). The lowercase symbols \( k \) and \( \hat{k} \) will be reserved for \( K \) and \( \hat{K} \) without the multilinear entries, i.e. for \( N = 2 \), \( k = [k_1, k_2] \) and \( \hat{k} = [\hat{k}_1, \hat{k}_2] \).

Suppose for the 'true' value of \( k \) that both \( P(s,k) \) and \( Q(s,k) \) are coprime. Then there exists the following (observability form) minimal state variable realization of (2):

\[
\frac{d}{dt} x_p(t) = [F + g_1(K)e^T]x_p(t) + g_2(K)u
\]

\[
y(t) = e^T x_p(t)
\]

(3)
where \( f(s) = \det(sI - F) \) is Hurwitz, \( e = [1, 0, \ldots, 0]^T \),

\[
\begin{align*}
g_1(K) &= G_1K + g_{10} \\
g_2(K) &= G_2K + g_{20}
\end{align*}
\]

with \( E \) the \( n \times (n-1) \) matrix \([I_{n-1}, 0]^T\) and \( f = [f_1, \ldots, f_n]^T\),

\[
F = [-f, E]
\]

and \( G_1, \ g_{10}, \ G_2 \) and \( g_{20} \) are defined as follows. If \( G_i \) has columns \( g_{i0} \), then with \( f(s) \) the \( n \)-th order monic polynomial with the coefficients in descending powers of \( s, f_1, \ldots, f_n \) respectively,

\[
\begin{align*}
f(s) - [s^{n-1}, \ldots, 1]g_{10} &= q_0(s) \\
[s^{n-1}, \ldots, 1]g_{20} &= p_0(s) \\
[s^{n-1}, \ldots, 1]g_{10} &= q_p(s) \\
[s^{n-1}, \ldots, 1]g_{20} &= p_p(s)
\end{align*}
\]

In the above the \( f_i \) are arbitrary, subject to \( f(s) \) being Hurwitz. The choice of the \( f_i \) of course fixes \( G_1 \) and \( g_{10} \). The controller will be of the form

\[
\begin{align*}
\frac{dx(t)}{dt} &= C_1(\hat{K}(t))x(t) + C_2(\hat{K}(t))y(t) + C_3(\hat{K}(t))r(t) \\
u(t) &= C_4(\hat{K}(t))x(t) + C_5(\hat{K}(t))y(t) + C_6(\hat{K}(t))r(t)
\end{align*}
\]

When the parameter \( k \) is perfectly known, the controller will be defined to achieve some objective such as pole positioning, LQG optimality, etc. We now impose the following assumptions on the plant and the controller.

**Assumption 1**

The closed-loop system is stable, with no unstable pole-zero cancellations when \( \hat{k} = k \). Here the elements of \( \hat{k} \) are the first \( N \) elements of \( \hat{K} \).

**Assumption 2**

The functions \( C_i \) are piecewise continuous and finite for finite \( \hat{k} \). Moreover, there exists an \( \epsilon_1 > 0 \) such that they are Lipschitz if \( \| \hat{k} - k \| < \epsilon_1 \).

**Assumption 3**

There exists an \( m_1 > 0 \) such that \( \| C_3(\hat{K}) \| \| C_2(\hat{K}) \| \| C_1(\hat{K}) \| + \| C_5(\hat{K}) \| C_6(\hat{K}) \| \neq 0 \) for all \( \hat{k} \) except on sets of measure zero. The inequality must also hold at \( \hat{k} = 0 \).

**Remark 1**

The first two assumptions ensure that the frozen closed-loop system is stable at \( \hat{k} = k \) and in a non-trivial region around it. The third assumption ensures that the frozen closed-loop transfer functions from \( r(t) \) and \( y(t) \) to \( u(t) \) are non-zero at all but isolated values of the parameter estimate \( \hat{k} \). This assumption ensures that the control input is persistently exciting at almost all values of the parameter estimate. The need for this to be true for \( \hat{k} = 0 \) arises because the identifier below requires \( \hat{k}(0) = 0 \).
Identifier

In formulating the identifier, as is common to all continuous time adaptive identifiers, one must first implement state variable filtering to avoid the need for explicitly differentiating the system signals when formulating the underlying adjustable model. To this end we formulate the following equations:

\[ \frac{d}{dt} x_1(t) = F^T x_1(t) + ey(t) \]  
\[ \frac{d}{dt} x_2(t) = F^T x_2(t) + eu(t) \]  
\[ h_0(t) = y(t) - g_{10}^T x_1(t) - g_{20}^T x_2(t) \]  
\[ H(t) = G_{11}^T x_1(t) - G_{21}^T x_2(t) \]  
(14)  
(15)

As will be shown later, from (3)-(10) one has that the underlying input/output relationship typified by (2) is equivalently described by the representation

\[ h_0(t) + K^T H(t) = 0 \]

Equivalently, with

\[ R(t) = \int_0^\infty e^{-\alpha(t-\tau)} H(\tau) H^T(\tau) \, d\tau \]

and

\[ r_0(t) = \int_0^\infty e^{-\alpha(t-\tau)} h_0(\tau) H^T(\tau) \, d\tau \]

for some \( \alpha > 0 \) one obtains

\[ R(t) K = -r_0(t) \]

Observe that \( r_0(t) \) and \( R(t) \) admit the on-line implementation

\[ \frac{d}{dt} R(t) = -\alpha R(t) + H(t) H^T(t), \quad R(0) = 0; \quad t \leq t_0 \]  
\[ \frac{d}{dt} r_0(t) = -\alpha r_0(t) + H(t) h_0(t), \quad r_0(0) = 0 \]  
(17)

Notice that \( R(t) \) is a least-squares matrix with exponential data forgetting. Its role is to give greater weight to more recent data as compared with data acquired at more remote instants of time. Consequently, adaptive algorithms using such a weighted data matrix have the ability to better track time variations in system parameters.

As in Reference 9, we will implement a two-step identification procedure. The first step provides an unconstrained parameter estimate \( K_u \), obtained as

\[ K_u(t) = -R^{-1}(t) r_0(t); \quad t \geq t_0 \]

for some positive \( t_0 \). This step is implemented only after some finite non-zero time \( t_0 \) has elapsed, because it requires the invertibility of \( R(t) \). Clearly, \( R(t) \) as presented is not invertible at \( t = 0 \). The question of whether \( R(t) \) ever becomes non-singular is tied into a persistent excitation requirement to be addressed later. For the moment we note that the explicit matrix
inversion at each instant of time can be avoided through the alternative equivalent implementation

\[ \frac{d}{dt} M(t) = \alpha M(t) - M(t)H(t)H^T(t)M(t), \quad M(t_0) = R^{-1}(t_0); \quad t \geq t_0 \]  

(18)

\[ K_u(t) = -M(t)R_0(t) \]  

(19)

Observe that (16) need only be implemented until \( t = t_0 \). In general, the generation of the unconstrained estimate does not in any way exploit the prior knowledge inherent in the setting being considered. It thus disregards the fact that in \( K \) certain elements are products of some others. The second step of our two-step procedure seeks to incorporate these constraints and so generates a constrained estimate \( \hat{K}(t) \):

\[ \frac{d}{dt} \hat{K}(t) = \frac{\partial \hat{K}(t)}{\partial K(t)} \Lambda [\hat{K}(t) - K_u(t)]; \quad t \geq t_0; \quad \hat{K}(t) = 0 \quad \forall t \in [0, t_0) \]  

(20)

Here the designer \( \Lambda = \text{diag}\{\lambda_0\}_{\sigma < \sigma} \) with \( \lambda_0 > 0 \).

Given (18), it seems that (20) is unnecessary. This is indeed so under conditions of perfect modelling. However, in actual settings the modelling is never perfect and one is faced with inevitable departures from idealizing assumptions. Such departures could take the form of noise, non-linearities or unmodelled high-frequency dynamics. As such, under these non-ideal circumstances one cannot expect \( K_u \) to have the required constrained structure, e.g. in the two-parameter case its third element will not in general be the required product of the first two.

This is where the need for the second step of our algorithm arises. This second constraining step introduces a gradient descent algorithm which locates the value of \( K \) that satisfies the constraints and is in some sense the closest to \( K_u \). It can thus be viewed as an attempt to use the implicit non-linear relations to filter out the imperfections and to constrain the eventual estimate to lie in the correct model. Such greater utilization of \textit{a priori} information results in the amelioration of imperfections introduced in the first step. Simulations presented in Section 4 verify this fact.

\textit{Convergence proof}

The convergence proof given below assumes the non-singularity of the data matrix \( R(t) \). Its satisfaction is discussed at length in Section 3. For the time being we assume that \( R(t) \) is uniformly positive definite at and beyond \( t_0 \).

Theorem 1 below shows that under the assumption of non-singularity of \( R(t_0) \) the parameter estimates converge uniformly asymptotically and all closed-loop signals are stable. The proof requires the following lemma proved by Dasgupta et al.\(^9\)

\textit{Lemma 1}

Consider the differential equation

\[ \frac{d}{dt} \hat{K}(t) = \frac{\partial \hat{K}(t)}{\partial K(t)} \Lambda [\hat{K}(t) - K] \]

with all quantities defined as above and \( \hat{K}(0) = 0 \). Then \( \hat{K}(t) - k(t) \) converges uniformly asymptotically to zero. Moreover, the convergence is locally exponential.

We now state and prove Theorem 1.
Theorem 1

If there exists a $t_0 \in \mathbb{R}_+$ such that $R(t_0)$ is non-singular, then with bounded $r(t)$, for (2)–(20)
(i) $\hat{k}(t)$ converges uniformly asymptotically to $k$ and
(ii) the closed-loop state $[x_p^T, x_c^T]^T$ and hence all system signals are bounded.

Proof. (i) Equation (2) implies

$$\frac{q_0(s)y(t) - p_0(s)u(t)}{f(s)} + \sum_{\sigma \in \sigma} \left( \prod_{i \in \sigma} k_i \right) \frac{q_0(s)y(t) - p_0(s)u(t)}{f(s)} = 0$$

(21)

Notice from (14) that

$$h_0(t) = [1 - g_0^T(sI - F^T)^{-1}e]y(t) - [g_0^T(sI - F^T)^{-1}e]u(t)$$

whence from (7) and (8)

$$h_0(t) = \left(1 - \frac{f(s) - q_0(s)}{f(s)}\right)y(t) - \frac{p_0(s)}{f(s)} u(t) = \frac{q_0(s)}{f(s)} y(t) - \frac{p_0(s)}{f(s)} u(t)$$

(22)

Similarly, (9), (10) and (15) yield

$$h_0(t) = \frac{q_0(s)}{f(s)} y(t) - \frac{p_0(s)}{f(s)} u(t)$$

(23)

where $h_0(t)$ are the elements of $H(t)$. Thus indeed (21) can be written as

$$h_0(t) + K^TH(t) = 0$$

(24)

whence, as noted earlier,

$$R(t)K + r_0(t) = 0$$

(25)

Thus from (19)

$$K_u(t_0) = K$$

(26)

Now, using (17)–(19), we have

$$\frac{d}{dt} K_u(t) = -M(t) \frac{d}{dt} r_0(t) - \frac{d}{dt} (M(t)) r_0(t) = -M(t)H(t) [h_0(t) + K_u^T(t)H(t)]$$

(27)

From (24)–(26) the term within the square brackets is zero at $t_0$. Thus $K_u(t) = K$ for all $t \geq t_0$. Hence with $\hat{k}(t_0) = 0$,

$$\frac{d}{dt} \hat{k}(t) = \frac{\partial K(t)}{\partial \hat{k}(t)} \Lambda [\hat{k}(t) - K] \quad \text{for all } t \geq t_0$$

(28)

Thus from Lemma 1 $\hat{k}(t)$ converges uniformly asymptotically to $k$.

(ii) Equations (3) and (11) combine to give for the appropriate matrices $A$, $b$ and $\Delta A$

$$\begin{bmatrix} \frac{d}{dt} x_p(t) \\ \frac{d}{dt} x_c(t) \end{bmatrix} = [A(k) + \Delta A(\hat{k}(t))] \begin{bmatrix} x_p(t) \\ x_c(t) \end{bmatrix} + b(\hat{k}(t)) r(t)$$

(29)

where $b(\hat{k}(t))$ is bounded and from (i) and Assumption 3 $\Delta A(\hat{k}(t)) \to 0$ and from Assumption 1 $A(k)$ is asymptotically stable. Thus (ii) follows from Reference 10.
3. PERSISTENCE OF EXCITATION

Central to the convergence result of Section 2 was the requirement that $R(t_0)$ be invertible. Conditions under which this happens are now discussed. To this end we make the following assumptions on the reference input $r(t)$ and the polynomials $p_\sigma$ and $q_\sigma$.

**Assumption 4**

The signal $r(t)$ lies in the set $\Omega_\Delta [0, \infty)$, for some $\Delta > 0$, defined as follows: there exists a countable, possibly empty set $C_\Delta = \{t_1, t_2, \ldots\}$ with $t_{i+1} - t_i \geq \Delta$ such that $r(t)$ and its derivative are continuous and bounded on $\mathbb{R}^+ - C_\Delta$ and possess finite one-sided limits at each $t_i$.

**Assumption 5**

The system transfer function (2) is such that there exists no collection of real scalars $\{\theta_\sigma\}$ for which

$$T_p(s) = \frac{\sum_{\sigma \subset S} \theta_\sigma p_\sigma(s)}{\sum_{\sigma \subset S} \theta_\sigma q_\sigma(s)}$$

**Assumption 6**

If there exists a collection of real scalars $\{\theta_\sigma\}$ such that

$$\sum_{\sigma \subset S} \theta_\sigma p_\sigma(s) = \sum_{\sigma \subset S} \theta_\sigma q_\sigma(s) = 0$$

then

$$\theta_\sigma = 0 \text{ for all } \sigma \subset S$$

**Remark 2**

Violation of either of the last two assumptions will imply that more than one unconstrained parameter estimate $K_u$ will satisfy the input/output relationship, thus making identifiability and persistence of excitation (PE) impossible to achieve. Also, they implicitly require that $n \geq 2^{N-1}$.

We now state a PE condition (Theorem 2) under which $R(t_0)$ is invertible at some $t_0 > 0$. By Theorem 1 this guarantees the stability of the adaptive law. The result relies on the time invariance of the closed-loop system for $t \leq t_0$. Theorem 2 shows that this same condition ensures the uniform positive definiteness of $R(t)$ and $M(t)$ even beyond $t = t_0$ as long as (i) of Theorem 1 holds. Such a property ensures robustness of the closed loop to small, unavoidable modelling inaccuracies. Theorem 2 follows in a straightforward fashion from techniques in References 11 and 12. Theorem 3 uses the fact that the non-singularity of $R(t_0)$ guarantees uniform asymptotic convergence of the parameter estimates and the boundedness of system signals. It then follows from techniques in References 12 and 13. Thus both proofs are omitted.
Theorem 2

Consider the closed-loop system (2)-(20) with $\dot{k}(t)$ a constant (note that this holds for $t < t_0$). Assume that Assumptions 4–6 hold and there exist $\alpha_1$, $\alpha_2$, $T > 0$ such that for all $t \in \mathbb{R}_+$

$$\alpha_1 I \leq \int_t^{t+T} \eta(\tau)\eta^T(\tau) \, d\tau \leq \alpha_2 I \quad (30)$$

Here, with $n$ the degree of $Q(s, k)$ and $m$ the highest degree among the polynomials $p_0q_\beta - q_0p_\sigma$ and $p_\beta q_\sigma - q_\beta p_\sigma$ for all $\beta$, $\sigma \in S$,

$$\eta(t) = \left[ r(t), \frac{r(t)}{s + \gamma}, \ldots, \frac{r(t)}{(s + \gamma)^{2n+m-1}} \right]^T \quad (31)$$

with $\gamma > 0$. Then there exists a $\alpha_3 > 0$ such that for some $t_0 > 0$

$$R(t_0) > \alpha_3 I \quad (32)$$

Theorem 3

Under the conditions of Theorem 2 there exists a positive real $\alpha_3$ such that $R(t) > \alpha_3 I$ for all $t \geq t_0$.

4. SIMULATION RESULTS

In this section we present simulation results using adaptive controllers typified by our first algorithm. Suppose that the unknown system is the circuit in Figure 1 containing a current-controlled voltage source. Suppose that the system input is the output of the independent voltage source (denoted $V(t)$) and the output is the voltage across the capacitor and that the only unknown values in the system are the capacitor and inductor values. Designating their respective reciprocals to be $k_1$ and $k_2$, the following is the transfer function of the system:

$$T_p(s) = \frac{k_1 s + k_2}{2s^2 + (k_1 - k_2)s + k_1 k_2}$$

Assume that the ‘true’ values are $k_1 = 2$ and $k_2 = 3$. Clearly, the ‘true’ system is therefore unstable. We design a standard pole placement control law whose objective is to place the

![Figure 1. The unknown system](image-url)
closed-loop poles at $-1$. The controller parameters are held constant over the region where the condition number of the Sylvester matrix$^{14}$ corresponding to the 'estimated' plant exceeds 100. The state variable filters used have the transfer function $\frac{1}{(s + 1)^2}$, i.e. $f(s) = (s + 1)^2$. The identifier (20) has $t_0 = 2$.

Figures 2–4 illustrate the values of $\hat{k}_1$, $\hat{k}_2$ and $y$ respectively in the absence of modelling inadequacies. Observe that the output after $t = 3$ is quickly stabilized. Further, the parameter estimates converge rapidly to their desired values.

Next consider the presence of an unmodelled mode at $s = -60$. Figures 5–7 show $\hat{k}_1$, $\hat{k}_2$ and $y$ respectively. The output is stabilized despite the unmodelled mode. Figures 8 and 9 depict the first two elements of the unconstrained estimate respectively.

Observe that the unmodelled dynamics causes non-terminating fluctuations in both the constrained and unconstrained estimates; however, those in the latter are much more pronounced. Further, the constrained estimates are on the whole closer to the 'true' values of $k_1$ and $k_2$ than are their unconstrained counterparts. These facts are illustrated by the following statistics, which reflect the ranges of oscillations exhibited by the $\hat{k}_i$ and $k_{ui}$ subsequent to $t = 20$:

$$
\begin{align*}
\hat{k}_{1\max} &= 2.0136, & k_{u1\max} &= 1.9468 \\
\hat{k}_{1\min} &= 2.0034, & k_{u1\min} &= 1.8830 \\
\hat{k}_{2\max} &= 3.0277, & k_{u2\max} &= 2.9822 \\
\hat{k}_{2\min} &= 2.9771, & k_{u2\min} &= 2.8381
\end{align*}
$$

Thus the utilization of more $a$ priori knowledge results in improved performance.

![Figure 2. Value of $k_1$ with no undermodelling](image)
Figure 3. Value of $k_2$ with no undermodelling

Figure 4. Value of output $y$ with no undermodelling
Figure 5. Value of $k_1$ with undermodelling

Figure 6. Value of $k_2$ with undermodelling
Figure 7. Value of output $y$ with undermodelling

Figure 8. Unconstrained estimate of parameter $1$ with undermodelling
5. CONCLUSIONS

We have presented a two-step algorithm for the adaptive control of systems where the unknown parameters are a class of physical element values. The algorithm has been shown to be globally, uniformly asymptotically stable and capable of withstanding non-ideal situations.

REFERENCES