March 1988 LIDS-P-1759

A Class of Adaptive Controllers with Application to Robust Adaptive Control

Mohammed Dahleh
Dept. of Electrical Engineering
Texas A & M University
College Station, TX 77843

Munther Dahleh
Dept. of EE & Comp. Sci, and LIDS
Massachusetts Institute of Technology
Cambridge, MA 02139

March 30, 1988

Abstract

This paper deals with discrete time, single-input single-output systems. A characterization of a general class of adaptive controllers is given. Using this characterization it is shown that an adaptive controller based on the l^1 optimal design strategy leads to a globally convergent adaptive scheme. The use of adaptive l^1 controllers is motivated by the problem of robust adaptive control, and the results of this paper offer an alternative way of handling this problem.

A Class of Adaptive Controllers with Application to Robust Adaptive Control

Mohammed Dahleh
Dept. of Electrical Engineering
Texas A & M University
College Station, TX 77843

Munther Dahleh
Dept. of EE & Comp. Sci, and LIDS
Massachusetts Institute of Technology
Cambridge, MA 02139

1. Introduction

The main objective of adaptive control is to compensate for the lack of knowledge of the plant by having time-varying controllers whose parameters change according to some learning mechanism. For example, a combined estimation/control procedure has exactly that property, where the learning is achieved via an appropriate estimation algorithm. The goal of this paper is to present a new approach to the problem of adaptive control. The main objective is to derive conditions for the stability of a general adaptive control scheme, with the least dependence on the specific design procedure. Such an approach will serve as a characterization of a large class of adaptive controllers, and thus make it possible to impose design specifications. Investigating the input-output properties of these schemes makes it possible to choose from a class of adaptive controllers the one that is most suitable from a practical point of view. For example, if a reduction in the sensitivity to external disturbances and high frequency unmodeled dynamics is desired, a satisfactory controller will most likely be based on the H^∞ or l^1 design procedures.

This general approach is intended to supply a method of analysis, which can be used to investigate the stability of adaptive schemes regardless of the design method. This approach reduces the problem to that of investigating the exponential stability of a sequence of time-invariant operators, the continuity of the design procedure, and the rate of variation of the controller parameters and the estimates of the plant parameters. With this approach the stability of the typical adaptive controllers based on pole-placement, or minimum prediction error controllers, can be easily inferred. An illustration of this approach will be given by applying it to the adaptive scheme based on the l^1-design methodology.

In this paper, we will consider an indirect scheme based on the l^1-optimal design methodology. This approach is investigated in the case where the unknown plant is Single-Input Single-Output and admits a unique l^1-optimal controller. Our goal is to explore such a methodology and see whether it leads to a globally convergent adaptive scheme.

The l^1 optimal control problem is a design philosophy where a linear controller is designed to achieve both internal stability and optimal disturbance rejection. This approach may lead to better ways of dealing with disturbances in the presence of adaptation. In adaptive control the problem of dealing with disturbances is of extreme importance. In the literature it has been demonstrated [1] that the presence of bounded and persistent disturbances may lead to the failure of the adaptive scheme through the degradation of performance or even the loss of stability. This problem has been the target of intensive research (e.g [2,3]) under the banner of robust adaptive control. The approach used in tackling these problems is to analyze an adaptive scheme which is designed without taking disturbances into account, and investigate its robustness to external disturbances. The idea of using an l^1-optimal design procedure may turn out to be the right way of handling such a problem. The reason is that l^1-optimal controllers are constructed to achieve optimal rejection of disturbances.

The adaptive l^1 controller will be based on certainty-equivalence approach, where at each step the system parameters are estimated and the controller is implemented using the estimated parameters. At each estimation stage a modeling error is commited which affects the output of the estimated plant. The l^1 optimal controller will be constructed to minimize the effect of this error on the output. In this paper we will show that under reasonable assumptions an adaptive controller can be based on the l^1 design strategy, and the result is a globally convergent adaptive scheme.

2. Model, A Priori Information, Parameter Estimation

In this paper, we will concentrate on single-input single-output discrete-time systems. Such systems can be represented by the following transfer function

$$P = \frac{B(z)}{A(z)}$$

where B and A are polynomials in z, given by

$$A(z) = 1 + a_1 z + \cdots\cdots + a_{m_1} z^{m_1}$$

$$B(z) = b_1 z + \cdots\cdots + b_{m_2} z^{m_2}$$

The variable z represents the unit shift operator defined by $zr(t) = r(t-1)$ for any sequence $r(t)$, where t is an integer. The coefficients of B and A are not known a priori. However, we will assume knowledge of a bound on the degrees of A and B. This is included in the following assumption

(AS-1)

The integer $n = \max(m_1,\ m_2)$ is known a priori.

In the presence of output disturbances of the form $\frac{1}{A}d(t)$, the model can be written as follows

$$y(t) = \phi(t-1)^T \theta_0 + d(t) \tag{1}$$

where

$$\theta_0^T = [-a_1, \ldots, -a_n, b_1, \ldots, b_n]$$

$$\phi(t-1)^T = [y(t-1), \ldots, y(t-n), u(t-1), \ldots, u(t-n)]$$

$$|d(t)| \leq \Delta$$

The form (1) is the form usually used in connection with parameter estimation [4]. We will assume that Δ is known, and we will use the constrained projection algorithm with *dead zone* for the recursive estimation part of the scheme. This algorithm is described by

$$\hat{\theta}(t) = \text{Projection of } s(t) \text{ on } C$$

$$s(t) = \hat{\theta}(t-1) + \frac{\eta(t-1)\phi(t-1)}{c + \phi(t-1)^T\phi(t-1)}[y(t) - \phi(t-1)^T\hat{\theta}(t-1)]$$

$$\eta(t-1) = \begin{cases} \hat{\eta}(t-1), & \text{if } |y(t) - \phi(t-1)^T\hat{\theta}(t-1)| > 2\Delta; \\ 0, & \text{otherwise.} \end{cases}$$

with $\hat{\theta}(0)$ given, and $\hat{\eta}(t-1)$ is a positive sequence that satisfies

$$0 < \eta_{min} \leq \hat{\eta}(t) \leq \eta_{max} \leq 1$$

The choice of the estimation algorithm is not unique and other algorithms, for example least squares, will work equally well. At each instant of time the estimation algorithm supplies an estimate $\hat{\theta}(t)$ from which we obtain estimates of the polynomials A_t and B_t given by

$$A_t = 1 + \hat{a}_1(t)z + \cdots. + \hat{a}_n(t)z^n$$

$$B_t = \hat{b}_1(t)z + \cdots. + \hat{b}_n(t)z^n$$

Before we continue, with the description of the problem, we will list the following properties of the constrained projection algorithm that will be useful in the analysis [4]. Let $e(t) = y(t) - \phi(t-1)^T\hat{\theta}(t-1)$

(1) $\|\hat{\theta}(t) - \theta_0\| \leq \|\hat{\theta}(t-1) - \theta_0\| \leq \|\hat{\theta}(0) - \theta_0\|, \quad t \geq 1$

(2) $\lim_{N\to\infty} \sum_{t=1}^{N} \frac{\eta(t-1)(e(t)^2-4\Delta^2)}{c+\phi(t-1)^T\phi(t-1)} < \infty$. This implies

(a) $\lim_{t\to\infty} \frac{\eta(t-1)(e(t)^2-4\Delta^2)}{[c+\phi(t-1)^T\phi(t-1)]} = 0$

(b) $\limsup_{t\to\infty} \|\hat{\theta}(t) - \hat{\theta}(t-k)\| \leq \frac{2\Delta\sqrt{\eta_{max}}}{c}$ for any finite k

In the case of no disturbances, the above properties reduce to the usual estimation properties [4]. Even in this case, the above properties do not imply that the estimates converge to the true values. Moreover, the estimates may not converge at all. The properties of the estimation algorithm, however, imply that the estimates remain bounded, and their variation slows down as time progresses.

3. Characterization of a Class of Slowly Time-Varying Controllers

We will present a general approach of analyzing adaptive schemes from the point of view of slowly time-varying systems. In this approach the sequence of estimated plants is viewed as a slowly time-varying system. Controllers will be designed on the basis of the frozen-time parameters of the system, and thus form a sequence which again can be regarded as a time-varying controller. In this section we will give a theorem which insures, under suitable assumptions, the stability of the closed-loop system. This setup applies to the problem of indirect adaptive control, where the plant is estimated and on the basis of the estimates a controller is designed. The slow variation of the estimates of the plant parameters is an issue that will be discussed later on.

We will consider a sequence of plants $P_t = \frac{B_t}{A_t}$, where B_t and A_t are polynomials in the unit shift operator z. For notational purposes the following definition will be used

Notation: $A_t \in STV(T, \gamma)$ if A_t is slowly time-varying in the sense that there exists constants γ and T such that

$$\|A_t - A_\tau\|_{\mathcal{A}} \leq \gamma |t - \tau| \quad \text{for all} \quad t, \tau > T.$$

$\mathcal{A}$ is the space of functions whose elements are of the form $h(z) = \sum_{i=0}^{\infty} h_i z^i$ with $\{h_i\}$ an element of l^1, (i.e the space of l^∞-stable time-invariant operators). Due to the fact that stable functions are isomorphic to l^1 sequences, we will not distinguish in notation between an element in $\mathcal{A}$ or its isomorphic image in l^1. The above definition means that the rate of variation of the operator A_t can be controlled after a sufficiently large time period elapses. If the sequence of plants P_t was generated by an estimation procedure, such as in the case of adaptive control with no disturbances, then it follows that the numerator and denominator of P_t have the above slow variation property. At each instance of time a controller $C_t = \frac{M_t}{L_t}$ is designed on the basis of the frozen parameter plant P_t. Note that at a particular point in time P_t is a linear time invariant plant. Therefore the design of C_t may be accomplished by a variety of linear control design approaches. Typical design procedures used in the literature on adaptive control include, minimum prediction error, and pole-placement controllers. The approach presented here cover these cases and has the advantage of being applicable to more elaborate control techniques. The controller parameters satisfy the following identity for each time instance t

$$G_t = M_t B_t + A_t L_t$$

where G_t is the closed loop polynomial. The following general result will give a sufficient condition for the l^∞-stability of a class of adaptive controllers.

Theorem 1: Given that $P = \frac{B}{A}$ is a linear time invariant plant, and N is an integer such that the degrees of A and B are bounded by N. Denote by A_t and B_t the estimates of A and B at time t. Suppose a time-varying controller C is implemented with the following property

$$C(r-y) = u \text{ if and only if } M_t(r(t) - y(t)) = L_t u(t)$$

and

$$A_t L_t + M_t B_t = G_t$$

where L_t, M_t, and G_t are stable polynomials, and $\{r(t)\}$ is a bounded reference input. Suppose that the following conditions hold:

(a) The sequence of plants are slowly time-varying in the sense, $A_t \in STV(T_A, \gamma_A)$, $B_t \in STV(T_B, \gamma_B)$, for some $T_A,\ T_B\ < \infty$.

(b) The closed-loop polynomial and the sequence of controllers are slowly time-varying in the sense, $G_t \in STV(T_G, \gamma_G)$, $M_t \in GTV(T_M, \gamma_M)$, $L_t \in STV(T_L, \gamma_L)$ for some $T_G,\ T_M,\ T_L\ < \infty$.

(c) There exists an integer N_1 such that the degrees of L_t and M_t are bounded by N_1 for all time t.

(d) The zeros of G_t are contained in the complement of the disc $\{s \in C : |s| < 1 + \epsilon\}$, in the complex plane.

(e) The norms of the frozen parameter, linear time-invariant operators ${G_t}^{-1}$, L_t, M_t, are bounded uniformly in t.

Then there exists a non-zero constants γ such that if $\gamma_G,\ \gamma_B,\ \gamma_A,\ \gamma_M,\ \gamma_L\ \leq \gamma$, the time-varying compensator will result in a stable adaptive system.

Proof: The idea of the proof is very intuitive. The estimation scheme provides a sequence of estimates of the plant, which in turn, will be used to design the compensator. Viewing these estimates as a linear time-varying plant, the compensator is chosen to stabilize this plant by stabilizing its frozen time values. The conditions in the theorem are exactly the conditions needed to insure the stability of this fictitious system. The error signal $e(t)$ will appear as a disturbance of the above system, and hence the operator mapping it to the signals $u(t), y(t)$ is stable. Finally, property $2a$ of the estimation scheme insures boundedness of the error signal and consequently $u(t), y(t)$, resulting in a stable adaptive system. In the course of the proof we will use the following notation in manipulating time-varying polynomials.

$$A_tB_t = \sum_i \sum_j a_i(t)b_j(t)z^{i+j} = B_tA_t$$

and

$$A_t.B_t = \sum_i \sum_j a_i(t)b_j(t-i)z^{i+j} \neq B_t.A_t$$

for time-invariant polynomials $AB = A.B = B.A$. Also, we will use the notation $[A_t, B_t] = A_t.B_t - A_tB_t$.

The time-varying polynomials A_t and B_t are obtained from a parameter estimation algorithm, driven by the error term $e(t) = y(t) - \phi(t-1)^T\hat{\theta}(t-1)$. The following equations are the basic components of the adaptive scheme:

$$e(t) = A_{t-1}y(t) - B_{t-1}u(t) \tag{2}$$

$$L_tu(t) = M_t(-y(t) + r(t)) \tag{3}$$

$$G_t = M_tB_t + A_tL_t. \tag{4}$$

We note that the output disturbances are implicitly included in the error signal. The basic idea is to relate the sequences $\{u(t)\}$ and $\{y(t)\}$ to the sequences $\{e(t)\}$ and $\{r(t)\}$, and show that the resulting operator is l^∞-stable. Using the above three equations (2)-(4) this can be easily done and the resulting equations can be written as:

$$\begin{bmatrix} G_t + X_t & S_t \\ W_t & G_t + Z_t \end{bmatrix} \begin{bmatrix} u(t) \\ y(t) \end{bmatrix} = \begin{bmatrix} w(t) - M_te(t) \\ z(t) + L_te(t) \end{bmatrix} \tag{5}$$

where

$$X_t = [B_t, M_t] + [L_t, A_{t-1}] + L_t(A_{t-1} - A_t)$$

$$S_t = [B_t, L_t] + [L_t, B_{t-1}] + (B_t - B_{t-1})L_{t-1}$$

$$W_t = -[A_t, M_t] + [M_t, A_{t-1}] + M_t(A_{t-1} - A_t)$$

$$Z_t = [A_t, L_t] + [M_t, B_{t-1}] + M_t(B_t - B_{t-1})$$

The filtered signals $w(t)$ and $z(t)$ are given by:

$$w(t) = A_t.M_t r(t)$$

$$z(t) = B_t.M_t r(t)$$

For any time $\tau \in Z^+$, we can factor the frozen time operator G_τ, evaluate the equations at $t = \tau$, and consider the evolution of the system as a function of τ. The equations can be written as

$$\begin{bmatrix} I + H_\tau(G_t - G_\tau) + H_\tau X_t & H_\tau S_t \\ H_\tau W_t & I + H_\tau(G_t - G_\tau) + H_\tau Z_t \end{bmatrix} \begin{bmatrix} u \\ y \end{bmatrix} (\tau)$$

$$= \begin{bmatrix} H_\tau w(\tau) - H_\tau M_t e(\tau) \\ H_\tau z(\tau) + H_\tau L_t e(\tau) \end{bmatrix}$$

where the operator H_τ is the inverse of the frozen parameter operator G_τ, whose invertibilty is insured by the assumptions in the theorem. Note that t in the above equation is only a parameter and does not represent time, since the equations evolve as a function of τ. As it stands this notation may seem confusing, but explicit formulas for the evolution of the operator equations will be given later which will clarify this point. Our objective is to show that if the perturbing operators involved in the above expression have the fading memory property then the above system is l^∞-stable. In specific, the fading memory property will be used to insure that the following operator

$$Q = \begin{bmatrix} I & 0 \\ 0 & I \end{bmatrix} + \begin{bmatrix} H_\tau(G_t - G_\tau) + H_\tau X_t & H_\tau S_t \\ H_\tau W_t & H_\tau(G_t - G_\tau) + H_\tau Z_t \end{bmatrix}$$

has an inverse which is l^∞-stable. The proof of this fact will be done in two steps. First we will show how to control the norms of the terms in

$$F = \begin{bmatrix} H_\tau(G_t - G_\tau) + H_\tau X_t & H_\tau S_t \\ H_\tau W_t & H_\tau(G_t - G_\tau) + H_\tau Z_t \end{bmatrix},$$

and second we will use the obtained estimates to prove the invertibility. Note that in the sequel we will use the symbol C for constants that may not be the same.

Step 1:

To perform the first task we will look at the operator $H_\tau(G_t - G_\tau)$, and show how we can estimate its norm. We start by giving a representation for the time-varying operator $F_1 = H_\tau(G_t - G_\tau)$. Let $y(t)$ be an l^∞ sequence, we can write

$$(G_t - G_\tau)(y)(t) = \sum_{i=0}^{t} (g_t(t-i) - g_\tau(t-i))y(i)$$

and

$$x(t) = H_\tau(G_t - G_\tau)(y)(t) = \sum_{k=0}^{t}\sum_{i=0}^{k} h_\tau(t-k)(g_k(k-i) - g_\tau(k-i))y(i),$$

where $\{h_\tau(k)\}$ is the pulse response for the time-invariant frozen operator H_τ. By changing the order of summation we get

$$x(t) = H_\tau(G_t - G_\tau)(y)(t) = \sum_{i=0}^{t}[\sum_{k=i}^{t} h_\tau(t-k)(g_k(k-i) - g_\tau(k-i))]y(i).$$

By setting $t = \tau$ in the above expression we get the following representation

$$x(\tau) = \sum_{i=0}^{\tau} f_1(\tau, i)y(i),$$

and the kernel of the mapping is given by

$$f_1(\tau, i) = \sum_{k=i}^{\tau} h_\tau(\tau - k)(g_k(k-i) - g_\tau(k-i)).$$

The slow variation will be used in order to show that the operator F_1 has fading memory. For that reason we will look at

$$\beta(\tau) = \sum_{i=0}^{\tau} |f_1(\tau, i)| \le \sum_{k=0}^{\tau} |h_\tau(\tau - k)| \sum_{i=0}^{k} |g_k(k-i) - g_\tau(k-i)|.$$

The operator $G_t \in STV(T, \gamma_G)$, which means that for a sufficiently large time T and $\tau, k > T$ we have

$$\sum_{i=0}^{k} |g_k(i) - g_\tau(i)| \leq \sum_{i=o}^{\infty} |g_k(i) - g_\tau(i)| = \|G_k - G_\tau\|_1 \leq \gamma_G |k - \tau|.$$

The fact that the time variation of G_t can be controlled for large enough time results in the following bound for $\beta(\tau)$

$$\beta(\tau) \leq \sum_{k=o}^{T} \|G_k - G_\tau\|_1 |h_\tau(\tau - k)| + \gamma_G \sum_{k=T+1}^{\tau} |k - \tau||h_\tau(\tau - k)| \quad \text{for} \quad \tau > T.$$

Consider the first term on the right hand side of the above inequality. Using the fact that $\|G_t\|$ is uniformly bounded by a constant, say B we get

$$\sum_{k=0}^{T} |h_\tau(\tau - k)| \|G_k - G_\tau\|_1 \leq 2B \sum_{k=\tau-T}^{\tau} |h_\tau(k)|.$$

Since H_τ is bounded uniformly in τ, it follows that for any given $\epsilon_1 > 0$ there exists a constant T_1 such that

$$\sup_{\tau > T_1} \sum_{k=\tau-T}^{\tau} |h_\tau(k)| < \epsilon_1.$$

The second term on the right of the above inequality can be rewritten as

$$\gamma_G \sum_{k=T+1}^{\tau} |k - \tau||h_\tau(\tau - k)| \leq \gamma_G \sum_{k=0}^{\tau-T-1} |h_\tau(k)| k \leq \gamma_G \sum_{k=0}^{\infty} k |h_\tau(k)|.$$

Using the fact that H_τ has a uniformly bounded norm and its poles are bounded away from a disc of radius $1 + \epsilon$, it follows that there exists a positive constant C such that

$$\gamma_G \sum_{k=T+1}^{\tau} |k - \tau||h_\tau(\tau - k)| \leq C\gamma_G$$

At this point we can choose γ_1 small enough so that $\epsilon_1 + \gamma_1 C$ can be made arbitrarily small. That is for any given $\epsilon > 0$, there exists an integer T_2 such that

$$\sup_{\tau > T_2} \beta(\tau) < \epsilon \tag{6}$$

The norms of the rest of the terms in F can be handled in a similar manner. For example the operator $H_\tau[B_t, M_t]$ can be estimated by

$$\|H_\tau[B_t, M_t]\|_1 \leq C\|[B_t, M_t]\|_1$$

Controlling the term $[B_t, M_t]$ follows in a similar way the analysis given above for the term $H_\tau(G_t - G_\tau)$, and a similar conclusion follows.

Step 2:

The invertibility of the operator $I + F$ where F is give by

$$F = \begin{bmatrix} H_\tau(G_t - G_\tau) + H_\tau X_t & H_\tau S_t \\ H_\tau W_t & H_\tau(G_t - G_\tau) + H_\tau Z_t \end{bmatrix},$$

is in essence concerned with the solvability of the equation

$$y(t) + Fy(t) = e(t).$$

for $e(t) \in {l_2}^\infty$. Let $f(t,s)$ be the kernel representing the operator F. From the analysis presented in the first step it follows that there exists an integer T such that

$$M_1 = \sup_{\tau > T} \sum_{i=0}^{\tau} \|f(\tau, i)\| < 1$$

First we investigate the operator $I + F$ on the time segment $[0, T]$. On this time segment the operator $I + F$ is finite dimensional and is given by

$$\begin{pmatrix} I + f(0,0) & 0 & \dots & 0 \\ f(1,0) & I + f(1,1) & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ f(T,0) & f(T,1) & \dots & I + f(T,T) \end{pmatrix} \begin{pmatrix} y(0) \\ y(1) \\ \dots \\ y(T) \end{pmatrix} = \begin{pmatrix} e(0) \\ e(1) \\ \dots \\ e(T) \end{pmatrix}$$

Denote by P_T the truncation operator which maps a sequence into its first T terms. The operator $P_T(I + F)P_T$ which maps $P_T({l_2}^\infty)$ into $P_T({l_2}^\infty)$ is invertible, since in our case $f(t,t) = 0$. Therefore there exists a constant C, such that

$$\|P_T y\|_\infty \leq C\|e\|_\infty \tag{7}$$

To complete the proof of the first step we should be able to bound the term $(I - P_T)y$ in terms of $e(t)$, arising from the solution of $(I + F)y = e$. We have

$$(I - P_T)y + (I - P_T)Fy = (I - P_T)e,$$

which implies that

$$\|(I - P_T)y\|_\infty - \|(I - P_T)Fy\|_\infty \leq \|(I - P_T)e\|_\infty \leq \|e\|_\infty. \tag{8}$$

Investigate the term

$$(I - P_T)Fy(t) = \begin{cases} \sum_{i=0}^{t} f(t,i)y(i), & \text{if } t > T; \\ 0, & t \leq T. \end{cases}$$

This implies that

$$\|(I - P_T)Fy\| \leq \sum_{i=0}^{t} \|f(t,i)\|\|y(i)\| \quad \text{for } t > T$$

$$= \sum_{i=0}^{T} \|f(t,i)\|\|y(i)\| + \sum_{i=T+1}^{t} \|f(t,i)\|\|y(i)\|$$

$$\leq \|P_T y\|_\infty \sup_{t>T} \sum_{i=0}^{T} \|f(t,i)\| + \|(I - P_T)y\|_\infty \sup_{t>T} \sum_{i=T+1}^{t} \|f(t,i)\|.$$

Now from equation (8), we have

$$\|(I - P_T)y\|_\infty \leq M_1 \|P_T y\|_\infty + M_1 \|(I - P_T)y\|_\infty + \|e\|_\infty$$

Which implies that

$$\|(I - P_T)y\|_\infty \leq \frac{1}{1 - M_1} M_1 C \|e\|_\infty + \frac{1}{1 - M_1} \|e\|_\infty. \tag{9}$$

Combining the results from (7)-(9) we get the following bound,

$$\|y\|_\infty \leq \|P_T y\|_\infty + \|(I - P_T)y\|_\infty \leq k_1 \|e\|_\infty$$

for some positive constant k_1.

Therefore, what we have established thus far is that in (5) the sequences $\{u(t),\ y(t)\}$ are bounded by the sequences $\{e(t),\ w(t),\ z(t)\}$. Equivalently,

$$||\phi(t)|| \leq K_1 + K_2 \max_{\tau \leq t} |e(\tau)|$$

Now using property (a) of the estimation algorithm and the generalized key technical lemma (Appendix) it follows that e, u, y are bounded functions. ∎

Discussion Of Theorem 1 We will consider the following two cases:
1) The case where there are no disturbances, i.e $\Delta = 0$. Condition (a) in Theorem 1 is immediately satisfied from property (b) of the estimation scheme. In fact, γ_A, γ_B are smaller than any positive γ for T large enough. If one designs the compensator to be uniformly continuous with respect to the coefficients in A_t and B_t, with the stability region in the complement of the disc of radius $\epsilon + 1$, conditions (b,d) will be satisfied. The boundedness conditions (c,e) are satisfied in many adaptive control designs where certain continuity requirements are met, and when $||G_t||$ does not approach zero. Hence, any frozen time control design methodology which stabilizes the estimates and at the same time is continuous with respect to these estimates will result in stabilizing the unknown system. We will demonstrate this idea by showing that designs based on the l^1 methodology for the disturbance rejection problem can be used as a basis for a stable adaptive scheme.
2) The case where $\Delta \neq 0$. Property (b) does not guarantee condition (a) in Theorem 1. This means that the speed of the estimation scheme has to be controlled, after some finite T. This can be done by choosing η_{max} to be small enough. Also, it is worth noting that the speed of the estimation scheme need not be controlled for all time but it has to be controlled for large enough time. The question of how small does $\hat{\eta}$ have to be is difficult to answer a priori. Certainly, the estimates derived in Theorem 1 give a very clear idea about the tradoffs involved, but this issue remains dependent on the control scheme employed. A primitive solution of this problem is to reduce the speed of estimation on line, till the signals in the loop become bounded, and then hold it at that speed. Once again, the controller can be designed to vary continuously with the estimates to insure the rest of the conditions in Theorem 1. Note that these results hold without any assumption of pesistence of excitation to force the parameters to converge. Their value is most obvious

in showing the limitations of adaptive control in the presence of disturbances. Also, this characterization has the advantage of providing us with a large class of stabilizing adaptive controllers, which makes it possible to satisfy performance specifications by choosing an appropriate one.

4. Application to Robust Adaptive Control

In this part the techniques presented in theorem 1 will be specialized to the problem of adaptive l^1 control. First, we will show how the l^1-optimal control problem arises naturally within the context of adaptive control. The idea is to design the controller to minimize the effect of disturbances and error signal on the output. Intuitively, this means that the estimation algorithm is less affected by the disturbance, and hence produces more realistic estimates. Equivalently, minimizing the effect of the error on the output suggests that the graphs of the plant and the frozen-time estimates, restricted on the domain of all possible signals generated by the control scheme, are close. Since the bounds on the disturbances and error signals are l^∞ bounds, the l^1 methodology is most adequate in this setting. Roughly, one can explain the procedure as follows: The input/output sequences are related through the true model by

$$Ay(t) = Bu(t)$$

Assume at time τ we have available the estimates A_τ, B_τ of A and B. Using this estimate we can write the input/output relation as

$$A_\tau y(t) = B_\tau u(t) + e(t)$$

where $\epsilon(t)$ is an error signal given by

$$e(t) = (A_\tau - A)y(t) + (B - B_\tau)u(t)$$

The control law $u(t)$ is given by

$$u(t) = C_\tau(-y(t) + r(t))$$

where C_τ is constructed to stabilize the plant $P_\tau = \frac{B_\tau}{A_\tau}$, and $r(t)$ is a bounded reference signal. With this control law the output of the plant is given by

$$y = P_\tau(1 + P_\tau C_\tau)^{-1} r + \frac{1}{A_\tau}(1 + P_\tau C_\tau)^{-1} e(t)$$

The first goal is to find a controller C_τ that internally stabilizes P_τ. Internal stability means [5] that every element of the following matrix is a stable transfer function

$$H(P_\tau, C_\tau) = \begin{bmatrix} \frac{1}{1+P_\tau C_\tau} & \frac{C_\tau}{1+P_\tau C_\tau} \\ \frac{-P_\tau}{1+P_\tau C_\tau} & \frac{1}{1+P_\tau C_\tau} \end{bmatrix}$$

The family of all compensators that stabilize P_τ, usually denoted by $S(P_\tau)$, is appropriately parametrized via the YJB parametrization [5]. If we think of the error signal $e(t)$ as a bounded l^∞ sequence, then it is reasonable to choose a compensator C_τ from $S(P_\tau)$ that minimizes the induced operator norm. This is exactly an l^1-optimal control problem defined as follows

$$\inf_{C_\tau \in S(P_\tau)} \left\| \frac{1}{A_\tau}(1 + P_\tau C_\tau)^{-1} \right\|_{\mathcal{A}} \tag{10}$$

where

$$S(P_\tau) = \{C_\tau | C_\tau \text{ internally stabilizes } P_\tau\}$$

A more precise explanation of the above intuitive idea can also be furnished in the case where the plant estimates are in $STV(0,\gamma)$ with γ small enough, i.e the estimates are slowly varying with a rate γ for all time $t \geq 0$. From the previous section, if a controller satisfies the conditions in Theorem 1, we have $(r = 0)$:

$$\begin{bmatrix} u(\tau) \\ y(\tau) \end{bmatrix} = (I + F)^{-1} \begin{bmatrix} -H_\tau M_t \\ H_\tau L_t \end{bmatrix} e(\tau)$$

With γ small enough, it follows that $\|F\|$ is very small and the output is dominated by the term

$$y \simeq H_\tau L_t e(\tau) = H_\tau L_\tau e(\tau) + H_\tau [L_t, L_\tau] e(\tau)$$

where the last term is small since L_t is slowly varying. Hence

$$y \simeq H_\tau L_\tau e(\tau) = \frac{1}{A_\tau}(1 + P_\tau C_\tau)^{-1} e(\tau) = \psi_\tau e(\tau)$$

By minimizing the l^1 norm at each frozen-time, we minimize un upper bound on the worst case output since

$$\sup_\tau \sum_{k=0}^{\tau} |\psi_\tau(k)| \leq \sup_\tau \|\psi_\tau\|_1$$

With some extra effort, we can obtain accurate estimates of the above, however, this will take us far from the main theme, and hence will not be presented.

The above considerations suggest the following adaptive scheme. First the parameters are estimated and A_t and B_t are obtained. Based on these estimates the controller C_t is obtained by solving the above optimization problem (10). The controller $u(t) = C_t(-y(t)+r(t))$ is implemented until the next estimation point, and then the process is repeated.

The main difference between this approach and other adaptive schemes (e.g. adaptive pole-placement) is that at each stage the controller is chosen for its ability to reject the modeling error, as well as for its stabilizing properties. The controller can also be chosen to minimize other functions, for example the sensitivity or the complementary sensitivity functions. Such approaches can be handled in an analogous manner. In order to keep the presentation simple, and to avoid unnecessary technical difficulties, we will require further knowledge of the unknown plant. This is included in the following assumption

(AS-2)

> There exists a convex set C that contains the unknown plant parameters, such that every plant whose parameters lie in C has a unique l^1-optimal controller, has distinct unstable zeros, has no zeros on the unit circle, and has no unstable pole-zero cancellation. This set C is assumed to be known a priori.

Assumption (AS-2) means that the model parameters are known up to membership of the set C. This is again similar to the case of adaptive pole-placement, where it is necessary to know the model sufficiently closely to insure the continuous solvability of the pole-placement problem.

4.1 Review of l^1-optimal design

In this section, we will describe how to solve the l^1-optimal control problem. For more details on this problem see [6,7]. The problem is to find the compensator C that solves

$$\inf_{C \in S(P)} \left\| \frac{1}{A}(1+PC)^{-1} \right\|_{\mathcal{A}}$$

where

$$S(P) = \{C | C \text{ internally stabilizes } P\}$$

We start by giving a parametrization of $S(P)$, the set of all stabilizing compensators of $P = \frac{B}{A}$, where A and B are coprime. This parametrization can be achieved by using coprime factorization [5]. Let X and Y be two stable rational functions that satisfy the following Bezout's identity

$$BX + AY = 1 \tag{11}$$

The set of all stabilizing compensators is given by

$$S(P) = \left\{ C | C = \frac{X + AQ}{Y - BQ},\ Q \in \mathcal{A} \right\}$$

Using this parametrization the function to be minimized is given by

$$\frac{1}{A}(1+PC)^{-1} = Y - BQ \quad Q \in \mathcal{A}$$

Let $z_1, \cdots, z_s$ be the zeros of B that lie inside the open unit disc. Let $K = BQ$, then K can be any stable rational function that satisfies

$$K(z_i) = 0, \quad i = 1, \ldots, s$$

Therefore the optimization problem can be written as

$$\inf_{K \in \mathcal{A}, K(z_i) = 0} \|Y - K\|_{\mathcal{A}}$$

Now recall that $R \in \mathcal{A}$ if and only if $R = \sum_{i=0}^{\infty} r_i z^i$ and $\sum_{i=0}^{\infty} |r_i| < \infty$. The induced norm on any element of $\mathcal{A}$, say $\sum_{i=1}^{\infty} r_i z^i$, acting as an operator on bounded sequences through convolution is equal to the l^1 norm of the sequence $\{r_i\}$. Also $K(z_i) = 0$ if and only if $\sum_{j=0}^{\infty} k_j z_i^j = 0$, if and only if $< k, z_i^R >= 0$ and $< k, z_i^I >= 0$ where

$$z_i^R = \mathrm{Re}(1, z_i, z_i^2, \ldots..)$$

$$z_i^I = \mathrm{Im}(0, z_i, z_i^2, \ldots..)$$

The notation $<,>$ refers to the natural pairing between the spaces l^1 and l^∞. Define the set S as follows

$$S = \{K \in l^1 | < K, z_j^R >= 0 \text{and} < K, z_j^I >= 0, j = 1, .., s\}$$

The optimization problem becomes

$$\inf_K \|Y - K\|_1 \tag{12}$$

such that $K \in S$.

Problem (12) has been solved by using duality [6] and was shown to be equivalent to a finite dimensional optimization problem. Specifically,

$$\mu = \inf_{K \in S} \|Y - K\|_1 = \max_{\alpha_i} [\sum_{i=1}^{s} \alpha_i \mathrm{Re} Y(z_i) + \sum_{i=1}^{s} \alpha_{i+s} \mathrm{Im} Y(z_i)]$$

subject to

$$-1 \leq \sum_{i=1}^{s} \alpha_i \mathrm{Re}(z_i^j) + \sum_{i=1}^{s} \alpha_{i+s} \mathrm{Im}(z_i^j) \leq 1, j = 0, 1, 2, \cdots$$

By solving the dual problem the optimal value μ is calculated, and the optimal functional in the dual space will be given by

$$\tilde{r} = \{\tilde{r}_j\} = \left\{ \sum_{i=1}^{s} \alpha_i^* \mathrm{Re}(z_i^j) + \alpha_{i+s}^* \mathrm{Im}(z_i^j) \right\}$$

To construct the optimal solution K^* we use the alignment between $\psi = Y - K^*$ and $\tilde{r}$. This alignment condition is given by

$$< \psi, \tilde{r} >= \|\psi\|_1 \|\tilde{r}\|_\infty = \mu \tag{13}$$

The condition (13) is true if and only if

$$\sum_{i=0}^{\infty} \psi_i \tilde{r}_i = \|\psi\|_1 \|\tilde{r}\|_\infty$$

if and only if

$$\psi_i = 0 \text{ whenever } |\tilde{r}_i| \neq \|\tilde{r}\|_\infty \tag{14}$$

and

$$\psi_i \tilde{r}_i \geq 0 \tag{15}$$

$$\sum_{i=0}^{\infty} |\psi_i| = \mu \tag{16}$$

From the definition of $\tilde{r}$, it is clear that only a finite number of the $\tilde{r}_i$'s have unit magnitude, and $|\tilde{r}_i| < 1$ for all i 's greater than an integer N. This has the implication that the optimal point ψ will have a finite number of nonzero ψ_i. In addition, for ψ to be admissible $Y - \psi \in S$, which results in the following system of linear equations

$$\sum_{i=0}^{\infty} \psi_i z_j^i = Y(z_j) \tag{17}$$

Relations (14)-(17) characterize the l^1 -optimal solution.

At this point we would like to know what the controller looks like at each adaptation instance. At time t we have A_t and B_t from which we calculate Y_t and X_t using Bezout's identity (11). Then we solve an l^1 problem to obtain ${\psi_t}^* = Y_t - {K_t}^*$. The above analysis shows that ${\psi_t}^*$ will be a polynomial of a sufficiently large degree N. The compensator C_t is given by

$$C_t = \frac{1}{P_t} \frac{1 - A_t {\psi_t}^*(t)}{A_t {\psi_t}^*}$$

We want a coprime factorization for the compensator C_t. Recall that $1 - A_t{\psi_t}^* = B_t(X_t + A_t{Q_t}^*)$ is a polynomial, and expression $X_t + A_t{Q_t}^*$ is a stable rational function which is not necessarily a polynomial. Therefore, $X_t + A_t{Q_t}^*$ can cancel some stable zeros of B_t. Thus we can write

$$1 - A_t{\psi_t}^* = (B_t)_u (B_t)_{su} M_t$$

where we denote the stable polynomial that is canceled by $X_t + A_tQ_t{}^*$ by $(B_t)_{sc}$, the rest of the stable polynomial by $(B_t)_{su}$, the unstable part of B_t by $(B_t)_u$, and by M_t the numerator of $X_t + A_tQ_t{}^*$. Therefore, the compensator C_t is given by

$$C_t = \frac{M_t}{(B_t)_{sc}\psi_t{}^*}$$

Define the following polynomials

$$L_t = \sum_{k=0}^{n_L} \hat{l}_k(t)z^k = (B_t)_{sc}\psi_t{}^*$$

$$M_t = \sum_{k=0}^{n_M} \hat{m}_k(t)z^k = T_t$$

The closed loop polynomial is given by

$$M_tB_t + A_tL_t = (B_t)_{sc}$$

Implementing the control law

$$L_tu = M_t(-y + r)$$

results in

$$(B_t)_{sc}y = B_tM_tr$$

$$y = (B_t)_{su}(B_t)_uM_tr$$

4.2 Technical Lemmas

This section contains three technical lemmas, which establish the continuity of the l^1-optimal design as a function of the system parameters. The first lemma establishes that the degree of the polynomials M_t and L_t, which define the controller at each stage, is bounded. The second and third lemmas demonstrate the continuity of the minimum l^1 norm and the optimal solution with respect to the system parameters.

Lemma 1: There exists an integer N such that the degree of $\psi_t{}^*$ is less than or equal to N for all t sufficiently large.

Proof: In section 4.1, it was shown that the alignment condition implies that the optimal solution is a polynomial of a certain degree. To show that there is a uniform bound on the degrees of $\psi_t{}^*$ we start by exhibiting an a priori estimate of the degree of the optimal solution, which depends on the system parameters. Rewrite the constraints in the dual problem as

$$-1 \leq \tilde{r} = F\underline{\alpha} \leq 1 \tag{18}$$

where F is the following semi-infinite matrix

$$F = \begin{bmatrix} 1 & 1 \,.. & 1 & 0 & .. & 0 \\ \mathrm{Re}(z_1) & \ldots & \mathrm{Re}(z_s) & \mathrm{Im}(z_1) & .. & \mathrm{Im}(z_s) \\ \mathrm{Re}(z_1^2) & \ldots & \mathrm{Re}(z_s^2) & \mathrm{Im}(z_1^2) & .. & \mathrm{Im}(z_s^2) \\ . & \ldots & . & . & .. & . \\ . & \ldots & . & . & .. & . \end{bmatrix}$$

Note that the matrix F is written with the understanding that if z_j is real then the column $\mathrm{Im}(z_j^i)$ is deleted. Also, when z_j is complex then $\bar{z}_j$ is also present and thus the above matrix contains only $\mathrm{Re}(z_j^i)$ and $\mathrm{Im}(z_j^i)$. Let L be an integer, $L \geq 2s$. Denote by F_L the matrix comprised of the first L rows of F. Now F_L has a full row rank and thus has a left inverse, which satisfies

$$F_L^{\dagger} F_L = I$$

Therefore we can write

$$\underline{\alpha} = F_L^{\dagger} F_L \underline{\alpha}$$

$$\|\underline{\alpha}\|_1 = \|F_L^{\dagger} F_L \underline{\alpha}\|_1 \le \|F_L^{\dagger}\|_X \|F_L \underline{\alpha}\|_\infty$$

The norm $\| \ \|_X$ is the induced norm of $F_L^{\dagger}$ as an operator from $l_\infty^L \to l_1^L$. This norm is bounded above by the sum of the absolute values of the elements of $F_L^{\dagger}$. The fact that $\underline{\alpha}$ satisfies the constraints (18) means that $\|F_L \underline{\alpha}\|_\infty \le 1$, and thus we have the following bound on $\underline{\alpha}$

$$\|\underline{\alpha}\|_1 \le \|F_L^{\dagger}\|_X$$

Let $w = \max_i |z_i| < 1$. For any k we have

$$|\tilde{r}_k| < w^k \|\underline{\alpha}\|_1$$

also

$$|\tilde{r}_k| < w^N \|\underline{\alpha}\|_1 \text{ for all } k > N$$

$$|\tilde{r}_k| < w^N \|F_L^{\dagger}\|_X \text{ for all } k > N$$

If we choose N large enough so that $w^N \|F_L^{\dagger}\|_X \le 1$ it follows that

$$\|\tilde{r}\|_\infty = \max_{k \le N} |\tilde{r}_k|$$

As was outlined in section 4.1, the degree of ψ_t will be less than or equal to N. The key point of the proof is the a priori estimate on N defined by the inequality

$$w^N \|F_L^{\dagger}\|_X \le 1$$

By an application of Rouche's Theorem we can show that the location of the zeros of B_t that lie in the unit disc is a continuous function of the coefficients of B_t. Therefore, due to the fact that the coefficients of B_t and A_t are bounded, and the above iequality it follows that there exist integers N and T such that

$$\text{Degree}\psi_t{}^*(t) \le N \text{ for all } t > T \blacksquare$$

Since the following two lemmas involve continuity properties with respect to the plant parameters, we should clarify the sort of topology imposed on the set of plants. Each plant

$P = \frac{B}{A}$ corresponds to a point $(b_1, \ldots, b_n, a_1, \ldots, a_n)$ in the Euclidean space R^{2n}. The topology imposed on the set of all plants is the Euclidean topology imposed on the set of points corresponding to the coefficients.

Lemma 2: Define $\mu(P)$ by

$$\mu(P) = \inf_{C \in S(P)} \left\| \frac{1}{A}(1 + PC)^{-1} \right\|_1$$

$\mu(P)$ is a continuous function of P.

Proof: Let P_k be a sequence of plants converging to P. For each P_k, the minimum l^1 norm of $\frac{1}{A_k}(1 + P_k C)^{-1}$ will be denoted by μ_k. Given an $\epsilon > 0$, by virtue of the fact that P_k converges to P, and Y_k to Y due to the continuous solvability of Bezout's identity, there exists an integer N_1 such that B_k and B have the same number of zeros inside the open unit disc, and

$$|\mu_k - \mu'_k| < \epsilon \text{ for all } k > N_1 \tag{19}$$

where μ'_k is defined as the solution to the following optimization problem

$$\mu'_k = \max_{\underline{\alpha} \in S_k} \left[\sum_{i=1}^{s} \alpha_i \mathrm{Re}(Y(z_i)) + \sum_{i=1}^{s} \alpha_{i+s} \mathrm{Im}(Y(z_i)) \right]$$

The set S_k is given by

$$S_k = \left\{ \underline{\alpha} \middle| \left| \sum_{i=1}^{s} \alpha_i \mathrm{Re}(z_{ik}^j) + \alpha_{i+s} \mathrm{Im}(z_{ik}^j) \right| \leq 1 \right\}$$

where the $\{z_{ik}, i = 1, \cdots s\}$ are the zeros of B_k that lie inside the unit disc. Also by the convergence of P_k to P it is clear that there exists an integer N_2 such that for all $k > N_2$ we have the following set inclusion

$$S_k \subset S^\epsilon = \left\{ \underline{\alpha} \middle| \left| \sum_{i=1}^{s} \alpha_i \mathrm{Re}(z_i^j) + \alpha_{i+s} \mathrm{Im}(z_i^j) \right| \leq 1 + \epsilon \right\}$$

This implies that

$$\mu'_k \leq \max_{\underline{\alpha} \in S^\epsilon} \left[\sum_{i=1}^{s} \alpha_i \mathrm{Re}(Y(z_i)) + \sum_{i=1}^{s} \alpha_{i+s} \mathrm{Im}(Y(z_i)) \right]$$

$$= \max_{\|\tilde{r}\|_\infty \leq 1+\epsilon} < Y, \tilde{r} >= (1+\epsilon)\mu \tag{20}$$

Combining (19) and (20) we conclude that

$$\mu_k \leq \mu + \epsilon(\mu + 1) \text{ for all } k > \max(N_1, N_2) \tag{21}$$

With an exactly similar argument we can show that there exists an integer N_3 such that

$$\mu \leq \mu_k(1+\epsilon) + \epsilon \text{ for all } k > N_3 \tag{22}$$

The inequalities (21) and (22) imply the result if μ_k is a bounded sequence. The sequence μ_k is defined by

$$\mu_k = \max_{\underline{\alpha} \in S_k} \left[\sum_{i=1}^{s} \alpha_i \mathrm{Re}(Y_k(z_i)) + \sum_{i=1}^{s} \alpha_{i+s} \mathrm{Im}(Y_k(z_i)) \right] \tag{23}$$

As was shown in lemma (1), the fact that $\underline{\alpha}$ satisfies the set of constraints $-1 \leq F_k \alpha \leq 1$ imposes a bound on its l^1 norm. This bound was shown to be $\|\underline{\alpha}\|_1 \leq \|F_{kL}^\dagger\|_X$. Since P_k converges to P, for k sufficiently large $\|F_{kL}^\dagger\|_X$ is a bounded sequence, which implies that there exist a constant M such that $\|\underline{\alpha}\|_1 \leq M$. Taking the absolute value of (23) we get

$$|\mu_k| \leq M \max_i (|\mathrm{Re}(Y_k(z_i)), |\mathrm{Im}(Y_k(z_i))|)$$

The sequences $|\mathrm{Re}(Y_k(z_i))|$ and $|\mathrm{Im}(Y_k(z_i))|$ are bounded due to the continuity of the solution of Bezout's identity with respect to the system parameters [4]. Therefore, the sequence μ_k is bounded and the lemma follows. ∎

Lemma 3: Let $\psi(P)$ be the unique l^1 optimal solution that satisfies

$$\mu(P) = \|\psi(P)\|_1$$

$\psi(P)$ is a continuous function of P.

Proof: Let P_k be a sequence of plants converging to P. Corresponding to each P_k there exist an optimal l^1 solution denoted by ψ_k. By lemma (1) we know that there exist integers

M_1 and M such that for all $k > M_1$, the optimal l^1 solutions ψ_k are polynomials of degree at most M. Thus we have

$$\mu_k = \|\psi_k\|_1 = \sum_{i=0}^{M} |\psi_{ik}|$$

In lemma (2) we showed that μ_k is a bounded sequence, in the sense that there exist an integer N_1 and a constant R such that

$$|\mu_k| < R \text{ for all } k > N_1$$

which means that

$$\|\psi_k\|_1 = \sum_{i=0}^{M} |\psi_{ik}| < R \text{ for all } k > \max(N_1, M_1)$$

We can prove the convergence of ψ_k by showing that all its cluster points are the same. Let ψ be the optimal l^1 solution corresponding to P, and let ψ_{k_m} be a convergent subsequence of ψ_k. The existence of this subsequence is guaranteed by the compactness of the unit ball in R^M. Assume that

$$\lim_m \psi_{k_m} = \tilde{\psi}$$

Since the ψ_{k_m} 's are the l^1 optimal solutions for the plants P_{k_m}, they must satisfy the interpolation conditions

$$\psi_{k_m}(z_{ik_m}) = Y_{k_m}(z_{ik_m})$$

Where z_{ik_m} are the zeros of B_{k_m} that lie inside the open unit disc. Using the continuity of the solution of the Bezout's identity (11) with respect to the system parameters we obtain

$$\lim_m \psi_{k_m}(z_{ik_m}) = \lim_m Y_{k_m}(z_{ik_m})$$

$$\tilde{\psi}(z_i) = Y(z_i) \tag{24}$$

Also from lemma (2) we have

$$\lim_m \|\psi_{k_m}\| = \|\tilde{\psi}\| = \mu \tag{25}$$

Relations (24) and (25) mean that $\tilde{\psi}$ satisfies the interpolation conditions and achieves the minimum norm μ, therefore by the uniqueness assumption (AS-2) it follows

$$\tilde{\psi} = \psi$$

This means that all the cluster points of ψ_k are the same and equal to ψ. This proves the continuity of the optimal solution with respect to the system parameters. ∎

4.3 Convergence of the Adaptive Scheme

The previous results established the following facts:

(1) The estimates A_t and B_t remain bounded, and are slowly time-varying.

(2) The polynomials M_t and L_t are continuously dependent on A_t and B_t, and slowly time-varying.

(3) The coefficients of the polynomials M_t and L_t are uniformly bounded.

Theorem: Subject to the assumptions (AS 1-2) the above scheme leads to

(i) $\{u(t)\}$ is a bounded sequence.

(ii) $\{y(t)\}$ is a bounded sequence.

(iii) The closed-loop characteristic polynomial tends to $(B_t)_{sc}$ in the sense

$$\lim_{t\to\infty}[(B_t)_{sc}y(t) - G(t-1,z)r(t)] = 0$$

where

$$G(t-1,z) = \sum_{j=1}^{n} \hat{b}_j(t-1) \sum_{k=0}^{n_M} \hat{m}_k(t-1)z^{j+k}$$

Proof:

Look at the filtered signals

$$w(t) = A_t.(L_t u(t) + M_t y(t))$$

$$M_t e(t) = M_t.A_{t-1}y(t) - M_t.B_{t-1}u(t)$$

Therefore we can write $w(t)$ as

$$w(t) = (M_t B_t + A_t L_t)u(t) + [A_t.L_t - A_t L_t]u(t)$$

$$+[A_t.M_t - A_t M_t]y(t) + M_t e(t)$$

$$-[M_t.A_{t-1} - M_t A_t]y(t) + [M_t.B_{t-1} - M_t B_t]u(t)$$

Similarly

$$z(t) = (M_t B_t + A_t L_t)y(t) - L_t e(t)$$

$$+[B_t.L_t - B_t L_t]u(t) + [B_t.M_t - B_t M_t]y(t)$$

$$+[L_t.A_{t-1} - L_t A_t]y(t) - [L_t.B_{t-1} - L_t B_t]u(t)$$

The relation between $\{u(t), y(t)\}$ and $\{e(t), w(t), z(t)\}$ can be summarized by

$$U \begin{bmatrix} u(t) \\ y(t) \end{bmatrix} = \begin{bmatrix} w(t) - M_t e(t) \\ z(t) + L_t e(t) \end{bmatrix} \tag{26}$$

where U is given by

$$U = \begin{bmatrix} (B_t)_{sc} + [A_t.L_t - A_t L_t] + [M_t.B_{t-1} - M_t B_t] & [A_t.M_t - A_t M_t] - [M_t.A_{t-1} - M_t A_t] \\ [B_t.L_t - B_t L_t] - [L_t.B_{t-1} - L_t B_t] & (B_t)_{sc} + [B_t.M_t - B_t M_t] + [L_t.A_{t-1} - L_t A_t] \end{bmatrix}$$

(26) is a linear time-varying system with input $\{e(t), w(t), z(t)\}$ and outputs $\{u(t), y(t)\}$. The technical lemmas demonstrated the continuity of the l^1-optimal design. This and the fact that the estimates of the model parameters are bounded and slowly time-varying implied that the controller parameters are slowly time-varying. Therefore, the assumptions of Theorem 1 are satisfied, and the conclusion that the system is stable follows. Since $r(t)$, $w(t)$, and $z(t)$ are bounded signals, it follows that $\{u(t)\}$ and $\{y(t)\}$ (and hence $\phi(t)$) grow no faster than linearly in $e(t)$, that is

$$\|\phi(t)\| \leq C_1 + C_2 \max_{0 \leq \tau \leq t} |e(\tau)|$$

for some positive constants C_1 and C_2. Recall from the properties of the estimation algorithm that

$$\lim_{t \to \infty} \frac{e(t)^2}{1 + \phi(t-1)^T \phi(t-1)} = 0$$

Using the key technical lemma of [8], we conclude that

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

and $\{u(t)\}$, $\{y(t)\}$ are bounded.

Since $e(t) = A_{t-1}y(t) - B_{t-1}u(t)$, we can write

$$L_t e(t+1) = L_t.A_t y(t+1) - L_t.B_t u(t+1)$$

$$= [L_t.A_t - L_t A_t]y(t+1) - [L_t.B_t - L_t B_t]u(t+1)$$

$$+L_t A_t y(t+1) - B_t[M_t r(t+1) - M_t y(t+1)]$$

$$= [L_t.A_t - L_t A_t]y(t+1) - [L_t.B_t - L_t B_t]u(t+1)$$

$$+(B_t)_{sc} y(t+1) - B_t M_t r(t+1)$$

Taking the limit as $t \to \infty$ of both sides of the above expression, and using the boundedness of A_t, B_t, M_t, L_t, $\{y(t)\}$, $\{u(t)\}$, $\{r(t)\}$ we get

$$\lim_{t\to\infty}[(B_t)_{sc} y(t) - G(t-1, z)r(t)] = 0 \blacksquare$$

Acknowledgement: Munther Dahleh is supported by the army research office, Center for Intelligent Control, under grant DAAL03-86-K-0171.

Appendix

Lemma: (Generalization of the Key Technical Lemma [8]). Let the sequence $\{\eta(t-1)\}$ be given by

$$\eta(t-1)=\begin{cases}\hat{\eta}(t-1), & \text{if } e(t)^2>4\Delta^2;\\ 0, & \text{otherwise.}\end{cases}$$

where $\{\hat{\eta}(t)\}$ is a bounded sequence that satisfies $0<\eta_{min}\leq\hat{\eta}(t)\leq\eta_{max}\leq 1$. If the sequences $\{e(t)\}$ and $\{\phi(t)\}$ satisfy

(1) $\lim\limits_{t\to\infty}\frac{\eta(t-1)[e(t)^2-4\Delta^2]}{c+\|\phi(t-1)\|^2}=0.$

(2) Linear boundedness condition: There exists constants $0<c_1<\infty$, $0<c_2<\infty$ such that

$$\|\phi(t)\|\leq c_1+c_2\max_{0\leq\tau\leq t}|e(\tau)|$$

then it follows that

(i) $\{\|\phi(t)\|\}$ is bounded

(ii) $\limsup e(t)^2\leq\frac{\eta_{max}}{\eta_{min}}4\Delta^2$

Proof: The proof of this lemma is a straightforward generalization of the proof in [8], and its inclusion here is for the sake of completeness. First assume that $\{e(t)\}$ is bounded. The linear boundedness condition implies that $\{\phi(t)\}$ is also bounded, and by condition (2) it follows that

$$\limsup\eta(t-1)e(t)^2\leq\limsup\eta(t-1)4\Delta^2\leq\eta_{max}4\Delta^2$$

Consider the following expression which appears in the above limit

$$\eta(t-1)e(t)^2=\begin{cases}\hat{\eta}(t-1)e(t)^2, & \text{if } e(t)^2>4\Delta^2;\\ 0, & \text{otherwise.}\end{cases}$$

which implies that

$$\eta_{min}e(t)^2\leq\begin{cases}\eta(t-1)e(t)^2, & \text{if } e(t)^2>4\Delta^2;\\ \eta_{min}4\Delta^2, & \text{if } e(t)^2\leq 4\Delta^2.\end{cases}$$

From the above limsup expression we get

$$\eta_{min}\limsup e(t)^2\leq\max(\limsup\eta(t-1)e(t)^2,\eta_{min}4\Delta^2)$$

$$\leq \max(\eta_{max}4\Delta^2, \eta_{min}4\Delta^2) \leq \eta_{max}4\Delta^2$$

In other words

$$\limsup e(t)^2 \leq \frac{\eta_{max}}{\eta_{min}} 4\Delta^2$$

To complete the proof assume that $\{e(t)\}$ is unbounded. Let $\{t_n\}$ be a sequence of integers such that

$$|e(t)| \leq |e(t_n)| \quad \text{for all } t \leq t_n$$

$$\lim_n |e(t_n)| = \infty$$

and

$$e(t_n)^2 > 4\Delta^2$$

The linear boundedness assumption and the above constructed sequence imply that there exists constants $0 < k_1 < \infty$ and $0 < k_2 < \infty$ such that

$$c + \|\phi(t_n)\|^2 \leq k_1 + k_2 e(t_n)^2$$

Along the subsequence $\{t_n\}$ the estimation algorithm weighting factor $\eta(t_n - 1)$ is equal to $\hat{\eta}(t_n - 1)$ and therefore it satisfies

$$\eta_{min} \leq \eta(t_n - 1) \leq \eta_{max}$$

Using the above observations it follows that along the subsequence $\{t_n\}$ we have the following inequality

$$\frac{\eta_{min}(e(t_n)^2 - 4\Delta^2)}{k_1 + k_2 e(t_n)^2} \leq \frac{\eta(t_n - 1)(e(t_n)^2 - 4\Delta^2)}{c + \|\phi(t_n - 1)\|^2}$$

but

$$\lim_{t_n} \frac{\eta_{min}(e(t_n)^2 - 4\Delta^2)}{k_1 + k_2 e(t_n)^2} = \frac{\eta_{min}}{k_2} > 0$$

Therefore

$$0 < \frac{\eta_{min}}{k_2} \leq \lim_{t_n} \frac{\eta(t_n - 1)(e(t_n)^2 - 4\Delta^2)}{c + \|\phi(t_n - 1)\|^2}$$

which contradicts condition (1) in the theorem. Thus the assumption that $\{e(t)\}$ is unbounded is invalid and the conclusion of the theorem follows. ∎

REFERENCES

1. B. Egardt, *Stability of Adaptive Controllers*, Springer-Verlag, New York, 1979.
2. B. D. O. Anderson, etal al, *Stability of Adaptive Systems, Passivity and Averaging Analysis*, MIT Press, 1986.
3. Romeo Ortega, and Tang Yu, "Theoretical Results on Robustness of Direct Adaptive Controllers: A Survey," *Proc. IFAC*, vol. 10, pp. 1-15, Munich, July 1987.
4. Graham C. Goodwin and Kwai Sang Sin, *Adaptive Filtering Prediction and Control*, Prentice Hall, 1984.
5. M. Vidyasagar, *Control System Synthesis: A Factorizaton approach*, The MIT Press, 1985.
6. Munther A. Dahleh and J. Boyd Pearson, Jr., " l^1 -Optimal Controllers for MIMO Discrete-Time Systems," *IEEE Trans. Automat. Contr.*, vol. AC-32, no. 4,pp. 314-323, April, 1987.
7. Munther A. Dahleh and J. Boyd Pearson, Jr., " L^1 -Optimal Compensators for Continuous-Time Systems," *IEEE Trans. Automat. Contr.*, vol. AC-32, no. 10,pp. 889-895, October, 1987.
8. G.C. Goodwin, P.J. Ramadge, and P.E. Caines, "Discrete-time multivariable adaptive control," *IEEE Trans. Automat. Contr.*, vol. AC–25, no. 3, pp. 449–456.