A SEQUENTIAL OPTIMAL REDESIGN PROCEDURE FOR LINEAR FEEDBACK SYSTEMS

by

E. Polak, S. Salcudean and D. Q. Mayne

Memorandum No. UCB/ERL M85/15

28 February 1985

ELECTRONICS RESEARCH LABORATORY

College of Engineering
University of California, Berkeley
94720

A Sequential Optimal Redesign Procedure for Linear Feedback Systems

by

E. Polak*, S. Salcudean* and D. Q. Mayne**

*Department of Electrical Engineering
and Computer Science
University of California
Berkeley, Calif. 94720
USA

**Department of Electrical Engineering
Imperial College
London SW7-2BT
ENGLAND

Abstract

This paper presents a new approach to on-line control system tuning, based on worst case design using semi-infinite optimization, together with a plant uncertainty identification scheme which this approach requires.

1. Introduction

Semi-infinite optimization is emerging as a powerful tool in control system design [1, 2], making it possible to formulate and solve important classes of new design problems. In this paper, we show that semi-infinite optimization can be used in a novel and powerful way for on-line control system tuning.

In the simplest case, the designer is given plant identification information, either in the form of *nominal* plant parameter values or in the form of an *uncertainty set* containing the plant parameters, as well as a set of design specifications. The specifications are then transcribed into semi-infinite inequalities involving compensator coefficients, which form, in turn, the constraints of a semi-infinite optimal design problem (see, e.g. [1]). The solution of the optimal design problem is a vector of optimal compensator parameters.

Unfortunately, complex specifications may be inconsistent and hence the optimal design problem may have no solution. To overcome this difficulty, in this paper, we formulate optimal design in terms of a new concept: that of *achievable performance* which is expressed in terms of an unconstrained semi-infinite optimization problem which always has a solution.

A more complex design situation arises when an identifier is used to update plant parameter information while the closed loop system is in operation, exploiting the intuitively clear idea that one should be able to improve the actual control system performance as the plant identification improves. We show that, under certain conditions, a sequential redesign scheme, based on information updates, produces a linear time varying control system which is asymptotically stable. Furthermore, we show that this time varying control system has better properties when a plant parameter uncertainty set identifier is used, than when an asymptotic plant parameter identifier is used.

There are various schemes in the literature, see e.g. [3], for the asymptotic estimation of plant parameters. The notion of parametric uncertainty set identification is a recent by-product of the use of semi-infinite optimization in design [2] and has not been explored until now. Consequently, we present in this paper a scheme for the identification of a decreasing sequence of uncertainty sets for the parameters of a plant which is incorporated in a closed loop system, and we show that under certain hypotheses this uncertainty set shrinks to the true parameter values as time goes on.

2. Optimal Design via Performance Factors

Without loss of generality, consider the simple SISO sampled-data system in Fig. 1. Suppose that the continuous plant characterization is given only in the form of the z-transform transfer function

$$P^*(\tilde{\theta}, z) = \left[\sum_{k=0}^{n} \tilde{b}_k z^{-k}\right] \Big/ \left[\sum_{k=1}^{n} \tilde{a}_k z^{-k}\right] = \frac{n_p(\tilde{\theta}, z)}{d_p(\theta, z)}, \tag{2.1}$$

where we have used the notation $\tilde{\mathbf{a}} \triangleq [\tilde{a}_0 \; \tilde{a}_1 \cdots \tilde{a}_n]^T$, $\tilde{\mathbf{b}} \triangleq [\tilde{b}_1 \; \tilde{b}_2 \cdots \tilde{b}_n]^T$ and $\tilde{\theta} \triangleq [\tilde{a}^T \; \tilde{b}^T]^T \in \mathbb{R}^{2n+1}$. The compensator to be designed is also specified as a z-transform transfer function, to be implemented by means of a digital computer and a sample and hold:

$$C^*(\mathbf{x}, z) = \left[\sum_{k=0}^{m} d_k z^{-k}\right] \Big/ \left[\sum_{k=1}^{m} c_k z^{-k}\right] \triangleq \frac{n_c(\mathbf{x}, z)}{d_c(\mathbf{x}, z)}, \tag{2.2}$$

where $\mathbf{x} \in \mathbb{R}^q$ is a vector of *designable* coefficients of the compensator, e.g. the coefficients of the numerator. For the purpose of illustration, let us consider three typical performance requirements: (i) closed-loop stability, (ii) rise time, overshoot and settling time bounds on the step response and, (iii) disturbance rejection in a specified frequency range. In *nominal design*, the plant parameter vector $\tilde{\theta}$ is assumed to be known. In *worst case* design, the plant parameter vector is assumed to be known only to the extent that $\tilde{\Theta} \in A$, a compact uncertainty set in $\mathbb{R}^{2n+1}$. Since we wish to deal with both worst-case and nominal design, we include the parameter vector Θ in all our formulas.

First we consider closed loop stability. The characteristic polynomial of the closed loop system in Fig. 1 is given by $\chi(\mathbf{x}, \Theta, \lambda) \triangleq n_p(\Theta, \lambda) n_c(\mathbf{x}, \lambda) + d_p(\Theta, \lambda) d_c(\mathbf{x}, \lambda)$, where $\lambda \in \mathbb{C}$. Let $d(\lambda)$ be a polynomial of the same degree as χ, with all the zeros of $d(\lambda)$ contained in the open unit disc $D \triangleq \{ \lambda \in \mathbb{C} | \; |\lambda| < 1 \}$. Let $T(\mathbf{x}, \Theta, \lambda) \triangleq \chi(\mathbf{x}, \Theta, \lambda) / d(\lambda)$. Then, refering to the modified Nyquist stability test proposed in [1], we see that if the locus of $T(\mathbf{x}, \Theta, \lambda)$ traced out for $\lambda = \exp(j\omega)$, with $0 \le \omega < 2\pi$, stays out of a parabolic region in $\mathbb{C}$ containing the origin, then the closed loop system is exponentially stable, assuming that there are no hidden oscillations. The parabolic region has a boundary described by $v = b'_s u - b''_s$, where $b'_s, b''_s > 0$ and $s \in \mathbb{C}$ is given by $s = u + jv$. Hence the stability requirement, for the sampled-data system, leads to the semi-infinite inequality (in normalized form)

$$\frac{1}{b''_s} \{\mathrm{Im}[T(\mathbf{x}, \Theta, j\omega)] - b'_s \, \mathrm{Re}[T(\mathbf{x}, \Theta, j\omega)]^2 + 1\} \le 0, \tag{2.3}$$

$$\forall \; \Theta \in A, \forall \; \omega \in [0, 2\pi]$$

where, in the case of the nominal design, one sets $A = \{\tilde{\Theta}\}$. In the case of worst case design, A is a compact set. Note that the coefficients b'_s and b''_s are *negotiable* and that they can be used to control the conservatism of the stability test.

Since we are interested in the stability of the sampled-data system, not only viewed as a discrete time system, but also of the sampled-data system viewed as a continuous time system, we must make sure that there are no hidden oscillations. Referring to [9], we see that, assuming that the continuous time plant is described by a rational transfer function, these can be eliminated by choosing a sufficiently high sampling rate.

The step response requirements can be expressed in terms of upper and lower bounds over an appropiate time range. Thus, if we let $y(\mathbf{x}, \theta, t), t = 0,1,\ldots,$ denote the closed loop system unit step response from rest, the sampled peak overshoot requirement leads to a set of inequalities of the form

$$\frac{y(\mathbf{x}, \theta, t)}{b_o} - 1 \le 0, \quad \forall\, \theta \in A, \quad \forall\, t \in \{0, 1, \cdots, t_o\}, \tag{2.4a}$$

where $A = \{\tilde{\theta}\}$ in the case of nominal design. The sampled rise time requirement leads to a set of inequalities of the form

$$-\frac{y(\mathbf{x}, \theta, t)}{b_r} + 1 \le 0, \quad \forall\, \theta \in A, \quad \forall\, t \in \{t_r, t_{r+1}, \cdots, t_s\}, \tag{2.4b}$$

where, again, $A = \{\tilde{\theta}\}$ in the case of nominal design. The sampled settling time requirements lead to similar inequalities.

To complete our dscussion, we address the question of disturbance rejection over a frequency range. The problem is somewhat complicated by the fact that the disturbance is assumed to be defined for all time, while our z-transform transfer functions deal only with sampled signals. Let $d(t)$ be the continuous time output disturbance and let $y(\mathbf{x}, t)$ be the corresponding continuous time closed loop system output. Assuming that the sampling rate is preassigned, we can write

$$d(t) = d_s(t) + d_r(t) \tag{2.5a}$$

$$y(\mathbf{x}, t) = y_s(\mathbf{x}, t) + y_r(\mathbf{x}, t) \tag{2.5b}$$

where $d_s(t)$ and $y_s(\mathbf{x}, t)$ are the outputs of zero-order sample and hold circuits, with input $d(t)$ and $y(t)$, respectively. Then, it is easy to show that

$$\frac{\hat{y}_s(\mathbf{x}, s)}{d_s(s)} = H^{*}_{yd}(\mathbf{x}, e^{sh}) \tag{2.5c}$$

where $y_s(\mathbf{x}, s)$ and $d_s(s)$ are the Laplace transforms of $y_s(\mathbf{x}, t)$ and $d(s)$ respectively and $H^{*}_{yd}(\mathbf{x}, z)$ is the z-transform transfer funtion from d to y and h is the sampling period. We note that the sampled-data system has no effect on the reminder $d_r(t)$. The only way to make $d_r(t)$ and $y_r(t)$ small is to increase the sampling rate. Since the spectrum of $d_s(t)$ is very close to the spectrum of $d(t)$ when the sampling rate is sufficiently high (as it ought to be for the feedback system to be effective), when the output disturbance $d(t)$ is specified as having a Fourier spectrum concentrated in a frequency interval $[\omega_1 , \omega_2]$, the disturbance rejection requirement results in the following normalized inequality :

$$\frac{1}{b_d(\omega)} \mid H^{*}_{yd}(\mathbf{x}, \theta, e^{j\omega h}) \mid - 1 \leq 0 , \quad \forall\, \omega \in [\omega_1 , \omega_2] , \tag{2.5d}$$

where $H^{*}_{yd}(\mathbf{x}, \theta, z)$ is the transfer function from d to y, h is the sampling period and $b_d(\omega) > 0$ is a continuous bound function.

The above examples show that design specifications result in normalized inequalities of one of two possible forms:

$$\frac{\varphi_{1j}(\mathbf{x}, \theta, \nu)}{b_{1j}(\nu)} - 1 \leq 0 , \quad \forall\, \theta \in A , \forall\, \nu \in N_{1j} , \tag{2.6a}$$

$$- \frac{\varphi_{2j}(\mathbf{x}, \theta, \nu)}{b_{2j}(\nu)} + 1 \leq 0 , \quad \forall\, \theta \in A , \forall\, \nu \in N_{2j} , \tag{2.6b}$$

where all functions are continuous and the denominators are strictly positive over the range of interest. Such a system of inequalities can certainly be inconsistent. In the case of inconsistency, the designer may try to change the bound functions $b_{1j}(\nu)$, $b_{2j}(\nu)$. A better way, derived from multicriteria optimization considerations, see e.g. [6], is to use a vector of weights $\mathbf{w} = [\mathbf{w}_1^T\, \mathbf{w}_2^T]^T > 0$ and to define the *achievable performance factor* $p(A , \mathbf{w})$ by

$$p(A , \mathbf{w}) \triangleq \min_{(y, \mathbf{x}) \in (\mathbb{R}, \mathbb{R}^q)} \{ y \mid \; w_{1j}[\varphi_{1j}(\mathbf{x}, \theta, \nu) / b_{1j}(\nu) - 1] \leq y , \tag{2.7}$$

$$\forall\, \theta \in A , \forall\, \nu \in N_{1j} , \forall\, j \in J_1 ;$$

$$- w_{2j}[\varphi_{2j}(\mathbf{x}, \theta, \nu) / b_{2j}(\nu) + 1] \leq y ,$$

$$\forall\, \theta \in A , \forall\, \nu \in N_{2j} , \forall\, j \in J_2 ,$$

$$\underline{\mathbf{x}} \le \mathbf{x} \le \overline{\mathbf{x}} \,\}$$

where w_{1j} , w_{2j} are the components of $\mathbf{w}_1$, $\mathbf{w}_2$; J_1 , J_2 are index sets and the sets N_{1j} , N_{2j} are either intervals or finite sets. The vectors $\underline{\mathbf{x}}$, $\overline{\mathbf{x}}$ denote realizability bounds on the design parameters. As before, in the case of nominal design, $A = \{\widetilde{\Theta}\}$ in (2.7).

A more sophisticated, but also more complex way of using of weights $\mathbf{w}_j$ is to make them ν dependent functions. When this is done, weights can be used for response shaping and more subtle tradeoff than can be done with fixed weights.

Suppose that $p(A, \mathbf{w}) \le 0$. Then all the specification inequalities are satisfied. When $p(A, \mathbf{w}) > 0$, at least one specification inequality is violated and the designer may wish to increase its weight in (2.7) while decreasing other weights as a way of getting closer to desired performance. Thus, for a selected compensator structure, design tradeoffs are performed by repeatedly solving problem (2.7) (which always has a solution) with adjusted weights until satisfaction is achieved. If that does not take place, the compensator structure must be modified.

3. Identification and Sequential Redesign

We shall now explore the possibility of obtaining improved performance in a closed loop system, such as the one in Fig.1, by redesigning and updating the compensator whenever the identification of the plant transfer function $P^*(\widetilde{\Theta}, z)$ is improved in some sense. Although not required by the theory we are about to present, in practice such a scheme is likely to work better if the initial plant parameter estimate, or the initial uncertainty set $A(0)$ is such that the design vector $(y(0), \mathbf{x}(0))$ which minimizes (2.7) results in an exponentially stable closed loop system. We assume that a plant identifier is installed which produces either a sequence of plant parameter estimates $\{\Theta(t)\}$ or a monotonically decreasing sequence of plant parameter uncertainty sets $\{A(t)\}$ such that the true plant parameter $\tilde{\Theta} \in A(t)$ for all $t \in \mathbf{N}_+$, where $\mathbf{N}_+ \triangleq \{0,1,2,3,\ldots\}$. The following two theorems can be deduced from results on max functions in [7, 10].

Theorem 3.1: Let $\{\Theta(t)\}$ be a sequence of plant parameter estimates such that $\Theta(t) \to \tilde{\Theta}$ as $t \to \infty$, and let $\{(y(t), \mathbf{x}(t))\}$ be a sequence of optimal solutions to (2.7), corresponding to $A = \{\Theta(t)\}$ and a given weight vector $\mathbf{w}$. Then,

(i) $p(\{\Theta(t)\}, \mathbf{w}) \to p(\{\tilde{\Theta}\}, \mathbf{w})$ as $t \to \infty$.

(ii) any accumulation point $\hat{\mathbf{x}}$, of $\{\mathbf{x}(t)\}$, defines an optimal compensator for the actual system.

∎

We conclude from this theorem that if $p(\{\tilde{\Theta}\}, \mathbf{w}) < 0$,i.e. desirable performance can be achieved, then there exists a t_0 such that for all $t \geq t_0$, $p(\{\Theta(t)\}, \mathbf{w}) \leq 0$ and, furthermore, by continuity of the functions φ_{1j} , φ_{2j}, the performance inequalities (2.5a), (2.5b) are satisfied for $A = \{\Theta\}$, i.e. desirable performance is achieved for all $t \geq t_0$. A similar result holds true for the performance inequalities considered one at a time.

Theorem 3.2: Let $\{A(t)\}$ be a sequence of compact uncertainty sets in $\mathbb{R}^{2n+1}$, such that for all $t \in \mathbb{N}$, $\tilde{\Theta} \in A(t)$ and $A(t+1) \subset A(t)$ Next, let $\{(y(t), \mathbf{x}(t))\}$ be a sequence of optimal solutions to (2.7), corresponding to $A = A(t)$ and a given weight vector $\mathbf{w} > 0$. Then,

(i) $p(A(t+1), \mathbf{w}) \leq p(A(t), \mathbf{w})$ for all $t \in \mathbb{N}$ and $\{p(A(t), \mathbf{w})\}$ is bounded;

(ii) If $A_\infty = \bigcap_{t=0}^{\infty} A(t)$, and $p(A_\infty, \mathbf{w}) < 0$, then there exists a t_0 such that for all $t \geq t_0$, $p(A(t), \mathbf{w}) \leq 0$;

(iii) If $A_\infty = \{\tilde{\Theta}\}$, then $p(A(t), \mathbf{w}) \to p(\{\tilde{\Theta}\}, \mathbf{w})$ as $t \to \infty$ and any accumulation point $\hat{\mathbf{x}}$ of $\{\mathbf{x}(t)\}$ defines an optimal compensator for the actual system.

Since the sequence $\{p(A(t),\mathbf{w})\}$ is monotone decreasing, it follows that if there exists a t_0 such that $p(A(t_0),\mathbf{w}) \le 0$, then $p(A(t),\mathbf{w}) \le 0$ must hold for all $t \ge t_0$. Thus we see that there is one major difference between nominal and worst case control system sequential redesign. In the worst case sequential design, one can determine when the identification is good enough to ensure that desired performance requirements are satisfied, while in the case of nominal design such a determination does not appear possible.

Now suppose the conditions in Theorems 3.1 or 3.2 prevail and that each time the compensator $C(\mathbf{x},z)$ in Fig. 1 is redesigned, it is updated in the feedback system. Note that we can ensure by algorithmic means that the compensator vectors $\mathbf{x}(t)$, $t \in \mathbb{N}$, defined as in Theorem 3.1 or Theorem 3.2, converge to $\hat{\mathbf{x}}$, a local optimal compensator vector for the actual plant, for the given weight vector $\mathbf{w}$. The following theorem which assumes a state space realization for the feedback system in Fig. 1, shows that if the stability inequality (2.7) is satisfied for this $\hat{\mathbf{x}}$, then the time varying sampled-data system which results from the compensator updates is uniformly asymptotically stable (assuming that there are no hidden oscillations).

Theorem 3.3: Consider the linear time varying discrete time system

$$\xi(t+1) = \mathbf{G}(t)\,\xi(t)\quad ,t = 0,1,\cdots \tag{3.1}$$

where the $\mathbf{G}(t)$ are $N \times N$ matrices such that $\mathbf{G}(t) \to \hat{\mathbf{G}}$ as $t \to \infty$, where $\hat{\mathbf{G}}$ has all its eigenvalues into the open unit disc $D \in \mathbb{C}$. Then (3.1) is uniformly asymptotically stable.

The above theorem can be proved easily by constructing a Lyapunov function corresponding to the system

$$\xi(t+1) = \hat{\mathbf{G}}\,\xi(t)\quad ,t = 0,1,\cdots \tag{3.2}$$

and showing that this Lyapunov function is also a Lyapunov function for the system (3.1) for t sufficiently large.

4. A Scheme for Plant Parameter Uncertainty Identification

In the control literature there exist many schemes for estimating the parameters of a plant from input and disturbance corrupted output measurements [3], but it appears that there are none for obtaining families of uncertainty sets $\{A(t)\}$, as stipulated in Section 3. We shall describe a new uncertainty identification scheme and we shall give sufficient conditions on the command input and compensator updating procedure so that the resulting time varying control system is asymptotically stable.

The uncertainty identification scheme we are about to present is based on a simple observation. Consider the simplest plant described by

$$y(t+1) = b_1 u(t) \ , \quad t \in \mathbb{N}_+ \tag{4.1a}$$

$$z(t) = y(t) + d(t) \ , \ t \in \mathbb{N}_+ \tag{4.1b}$$

and suppose that the disturbance is known only to the extent that it is bounded, i.e.

$$|d(t)| \le \alpha \ , \ \forall t \in \mathbb{N}_+ \tag{4.2}$$

Now, suppose that we have measured $\{u(\tau)\}$ and $\{z(\tau)\}$ for $\tau = 0,1,\ldots,t$. Then, for all $\tau \le t$ such that $u(\tau-1) \ne 0$ we have that b_1 is given by

$$b_1 = \frac{y(\tau)}{u(\tau-1)} \tag{4.3}$$

but cannot be computed from (4.3) because $\{y(\tau)\}$ is not accessible. However, for all $t \in \mathbb{N}_+$ we can compute an interval $A(t)$ at time t which contains b_1 as follows. Note that from (4.3),

$$b_1 = \frac{z(\tau)-d(\tau)}{u(\tau-1)}, \quad \forall\, \tau \in \mathbb{N}_+ \text{ such that } u(\tau-1) \neq 0 \ . \tag{4.4}$$

Hence, since $d(\tau) \in [-\alpha, \alpha]$ for all $\tau \in \mathbb{N}_+$, we have that

$$b_1 \in A(t) \triangleq \bigcap_{\substack{\tau \leq t \\ u(\tau-1)\neq 0}} \left[\frac{z(\tau)-\alpha}{u(\tau-1)} \, , \, \frac{z(\tau)+\alpha}{u(\tau-1)} \right] . \tag{4.5}$$

Note that the interval $A(t)$ shrinks as t increases. Furthermore, suppose that for some $\tau_1, \tau_2 \in \mathbb{N}_+$ we have that $d(\tau_1) = \alpha$, $d(\tau_2) = -\alpha$ with $u(\tau_i-1) \neq 0$ for $i=1,2$. Then, by substituting (4.1b), (4.4) in (4.5) we obtain that

$$b_1 \in A(t) \triangleq \bigcap_{\substack{\tau \leq t \\ u(\tau-1)\neq 0}} \left[b_1 + \frac{d(\tau)-\alpha}{u(\tau-1)} \, , \, b_1 + \frac{d(\tau)+\alpha}{u(\tau-1)} \right] = \{b_1\} \tag{4.6}$$

Thus, for this smple case , it is possible to construct an identification scheme which

(i) produces a decreasing sequence of intervals containing the actual plant parameter , and

(ii) yields the actual parameter b_1 in finite time if the disturbance sequence attains both of its bounds in finite time.

We shall now show that this simple idea can also be extended to higher dimensional difference equations of known order.

Consider the n-th order plant defined in (2.1). Its input-output behavior is governed by the difference equation

$$\sum_{k=0}^{n} \tilde{a}_k \, y(t+n-k) = \sum_{k=1}^{n} \tilde{b}_k \, u(t+n-k) \, , \tag{4.7}$$

where the coefficients $\tilde{a}_k$ and $\tilde{b}_k$ need to be identified. We shall assume that we can measure the plant input $\{u(t)\}$ and disturbance corrupted output $\{z(t)\}$, where $z(t) = y(t)+d(t)$ for all $t \in \mathbb{N}_+$ and that there exists an $\alpha \in (0, \infty)$ such that the output disturbance satisfies $|d(t)| \leq \alpha$ for all $t \in \mathbb{N}_+$.

For a given plant input $\{u(t)\}$ and corresponding output $\{y(t)\}$, let

$$\mathbf{y}(t) \triangleq [y(t+n) \cdots y(t)]^T \in \mathbb{R}^{n+1}, \tag{4.8}$$
$$\mathbf{z}(t) \triangleq [z(t+n) \cdots z(t)]^T \in \mathbb{R}^{n+1},$$
$$\mathbf{d}(t) \triangleq [d(t+n) \cdots d(t)]^T \in \mathbb{R}^{n+1},$$
$$\mathbf{u}(t) \triangleq [u(t+n-1) \cdots u(t)]^T \in \mathbb{R}^{n},$$
$$\boldsymbol{\varphi}(t) \triangleq [y(t+n-1) \cdots y(t)\, u(t+n-1) \cdots u(t)]^T \in \mathbb{R}^{2n}.$$

Although it is common to normalize (4.7) by setting $\tilde{a}_0 = 1$, we will need the alternative normalization $\|\tilde{\mathbf{a}}\|_1 = 1$, where $\|\tilde{\mathbf{a}}\|_1 = \sum_{k=0}^{n} |\tilde{a}_k|$.

Definition 4.1: For a given plant input $\{u(t)\}$ and observed output $\{z(t)\}$, we define the ***plant uncertainty set*** $F(t)$, at time $t \geq n$ to be the set of all plant parameters $\theta = [\mathbf{a}^T\ \mathbf{b}^T]^T$, with $\mathbf{a} \in \mathbb{R}^{n+1}$, $\mathbf{b} \in \mathbb{R}^n$, and $\|\mathbf{a}\|_1 = 1$, such that the output $\{y(t)\}$ of the difference equation

$$\sum_{k=0}^{n} a_k\, y(t+n-k) = \sum_{k=1}^{n} b_k\, u(t+n-k) \tag{4.9}$$

satisfies $|y(\tau)-z(\tau)| \leq \alpha$ for all $\tau \leq t-n$, for all initial conditions $y(\tau), \tau = 0,1,\ldots,n-1$ such that $|y(\tau)-z(\tau)| \leq \alpha$ for $\tau = 0,1,\cdots n-1$, i.e., using the notation (4.8) and rewriting (4.9) as an inner product,

$$F(t) \triangleq \{\theta = [\mathbf{a}^T\, \mathbf{b}^T]^T \mid \mathbf{a} \in \mathbb{R}^{n+1},\ \|\mathbf{a}\|_1 = 1, \mathbf{b} \in \mathbb{R}^n, \tag{4.10}$$
$$\text{and there exists } d(\tau) \text{ such that } |d(\tau)| \leq \alpha,$$
$$\text{for } 0 \leq \tau \leq t, \text{ and}$$
$$\langle \mathbf{a}, \mathbf{z}(\tau) - \mathbf{d}(\tau) \rangle - \langle \mathbf{b}, \mathbf{u}(\tau) \rangle = 0 \quad \text{for } 0 \leq \tau \leq t-n \}.$$

∎

We note that, without additional assumptions, it is not possible to identify the plant at time t with any greater precision than to state that the actual parameter vector $\tilde{\theta} \in F(t)$. Since the description of $F(t)$ is too complex for efficient algorithmic design, we introduce a computationally more tractable uncertainty set $A(t)$ that contains the set $F(t)$.

Definition 4.2: For $t \in \mathbf{N}_+, t \geq n$ let

$$A(t) \triangleq \{ \Theta = [\mathbf{a}^T \ \mathbf{b}^T]^T \mid \mathbf{a} \in \mathbf{R}^{n+1}, \|\mathbf{a}\|_1 = 1, \mathbf{b} \in \mathbf{R}^n \text{ , and } |\langle \mathbf{a}, \mathbf{z}(\tau) \rangle - \langle \mathbf{b}, \mathbf{u}(\tau) \rangle| \leq \alpha \text{ for } 0 \leq \tau \leq t-n \} . \quad (4.11)$$

■

First, note that each inequality that appears in the definition of $A(t)$ determines a slab in $\mathbf{R}^{2n+1}$. Thus $A(t)$ is the intersection of a polyhedron, which is the intersection of slabs, and the surface defined by the piecewise linear constraint $\|\mathbf{a}\|_1 = 1$, see Fig.2 . Next, note that, as t increases, the number of intersecting slabs that form the polyhedron increases and therefore $A(t)$ shrinks as t increases. It is possible that the sets $A(t)$ shrink to the set $\{\tilde{\Theta}, -\tilde{\Theta}\}$ in finite time. To see how this may happen, suppose that $\tilde{\Theta} = [\mathbf{a}^T, \mathbf{b}^T]^T \in A(t)$. Then, from (4.11), for $\tau \leq t-n$, $|\langle \mathbf{a}, \mathbf{z}(\tau) \rangle - \langle \mathbf{b}, \mathbf{u}(\tau) \rangle| \leq \alpha$. Since $\langle \tilde{\mathbf{a}}, \mathbf{y}(\tau) \rangle - \langle \tilde{\mathbf{b}}, \mathbf{u}(\tau) \rangle = 0$ for all τ, we obtain that $-\alpha \leq \langle \mathbf{a}-\tilde{\mathbf{a}}, \mathbf{z}(\tau) \rangle - \langle \mathbf{b}-\tilde{\mathbf{b}}, \mathbf{u}(\tau) \rangle + \langle \tilde{\mathbf{a}}, \mathbf{d}(\tau) \rangle \leq \alpha$ must hold. Now, suppose that the same pair $(\mathbf{z}(\tau), \mathbf{u}(\tau)) \triangleq (\mathbf{p}, \mathbf{q})$ occurs for two values τ_1 and τ_2 of $\tau \in \{0, 1, \cdots t-n\}$, while $\langle \tilde{\mathbf{a}}, \mathbf{d}(\tau_1) \rangle = \alpha$ and $\langle \tilde{\mathbf{a}}, \mathbf{d}(\tau_2) \rangle = -\alpha$. Then the intersection of the slabs corresponding to $\tau = \tau_1$ and $\tau = \tau_2$ is a hyperplane orthogonal to $[\mathbf{p}^T \mid -\mathbf{q}^T]^T$ which contains $\tilde{\Theta}$. It follows that if this phenomenon occurs 2n times with the normals to the resulting hyperplanes being linearly independent, then the normalization $\|\mathbf{a}\|_1 = 1$ implies that $A(t)$ is reduced to $\{\tilde{\Theta}, -\tilde{\Theta}\}$ in finite time. In Theorem 4.1, we will show that under plausible assumptions on the disturbance sequence $\{d(t)\}$, the sets $A(t)$ shrink to $\{\tilde{\Theta}, -\tilde{\Theta}\}$.

Proposition 4.1: For all $t \geq n$, (i) $\tilde{\Theta} \in F(t) \subset A(t)$; (ii) $F(t+1) \subset F(t)$, and (iii) $A(t+1) \subset A(t)$.

Proof :

(i) Let $\Theta = [\mathbf{a}^T \mathbf{b}^T]^T \in F(t)$. Then, by (4.10), $\|\mathbf{a}\|_1 = 1$ and there exists $\{d(t)\}$ such that $|d(\tau)| \le \alpha$ for $0 \le \tau \le t$ and $\langle \mathbf{a}, \mathbf{z}(\tau) - \mathbf{d}(\tau)\rangle - \langle \mathbf{b}, \mathbf{u}(\tau)\rangle = 0$. It follows that, for all $0 \le \tau \le t-n$,

$$
\begin{aligned}
|\langle \mathbf{a}, \mathbf{z}(\tau)\rangle - \langle \mathbf{b}, \mathbf{u}(\tau)\rangle| &= |\langle \mathbf{a}, \mathbf{y}(\tau)\rangle - \langle \mathbf{b}, \mathbf{u}(\tau)\rangle + \langle \mathbf{a}, \mathbf{d}(\tau)\rangle| \\
&= |\langle \mathbf{a}, \mathbf{d}(\tau)\rangle| \le \alpha,
\end{aligned}
$$

and therefore $\Theta \in A(t)$. Since we have assumed that output disturbance is bounded by α and $\|\tilde{\mathbf{a}}\|_1 = 1$, it is clear that $\tilde{\Theta} \in F(t)$.

■

Definition 4.3: We shall say that a sequence $\{\xi(t)\} \subset \mathbb{R}^n$ is *persistently exciting* (p.e.) if there exists a t^* and a $\delta > 0$ such that for all $k \in \mathbb{N}_+$,

$$
\sum_{t=k}^{k+t^*} \xi(t)\, \xi(t)^T \ge \delta\, \mathbf{I} \ . \tag{4.12}
$$

■

Note that in the above definition, we do not require an upper bound on the sum of the rank one matrices $\xi(t)\,\xi(t)^T$ in (4.12) , so that the definition is not restricted to bounded sequences.

Consider the intersection $S \cap C_\rho(\tilde{\Theta})$ of the polyhedral surface $S \triangleq \{\Theta \in \mathbb{R}^{2n+1} \mid \Theta = [\mathbf{a}^T \mathbf{b}^T]^T, \|\mathbf{a}\|_1 = 1\}$ with the half cylinder of radius ρ, $C_\rho(\tilde{\Theta}) = \{\Theta \in \mathbb{R}^{2n+1} \mid \Theta = \mu\tilde{\Theta} + \mathbf{w}, \mu \ge 0, \langle \mathbf{w}, \tilde{\Theta}\rangle = 0, \|\mathbf{w}\|_2 \le \rho\}$. Clearly, the diameter of $S \cap C_\rho(\tilde{\Theta})$ is proportional to ρ. The following lemma establishes a bound on ρ such that for all $t \ge k^* + n$, , where k^* is as in (4.12) , $A(t) \subset S \cap (C_\rho(\tilde{\Theta}) \cup C_\rho(-\tilde{\Theta}))$ and hence on how well $A(t)$ approximates $\tilde{\Theta}$.

Definition 4.4 For all $t \in \mathbb{N}_+$ let $A^\pi(t)$ denote the *orthogonal projection* of $A(t)$ on the orthogonal complement of the line spanned by $\tilde{\Theta}$, i.e.

$$
A^\pi(t) \triangleq \{\mathbf{v} \in \mathbb{R}^{2n+1} \mid \langle \mathbf{v}, \Theta\rangle; \langle \Theta - \mathbf{v}, \mathbf{v}\rangle = 0 \text{ for some } \Theta \in \mathbb{R}^{2n+1}\} \tag{4.13}
$$

and let the $\|\cdot\|_2$-norm radius of $A^\pi(t)$ be defined by

$$\rho(t) \triangleq \sup \{ \|\mathbf{v}\|_2 \mid \mathbf{v} \in A^{\pi}(t) \} \tag{4.14}$$

Lemma 4.1: Suppose that the sequence $\{\varphi(t)\}$ defined in (4.8) is p.e., with $t^* \in \mathbb{N}_+$ and $\delta > 0$ such that for all $k \in \mathbb{N}_+$

$$\sum_{t=k}^{k+t^*} \varphi(t)\, \varphi(t)^T \geq \delta\, I \,. \tag{4.15}$$

Then, for all $t \geq t^*+n$, (i) $A(t)$ is bounded, and (ii) $\rho(t) \leq 2\alpha \sqrt{t^*+1} \,/\, \delta$.

Proof: For all $t \in \mathbb{N}_+$, let $\mathbf{G}(t) \in \mathbb{R}^{(t^*+1)\times(2n+1)}$ be defined by

$$\mathbf{G}(t) \triangleq \begin{bmatrix} \mathbf{y}(t)^T & -\mathbf{u}(t)^T \\ \cdot & \cdot \\ \cdot & \cdot \\ \cdot & \cdot \\ \mathbf{y}(t+t^*)^T & -\mathbf{u}(t+t^*)^T \end{bmatrix} \tag{4.16a}$$

$$\triangleq \left[\begin{array}{c|c} y(t+n) & \cdot \\ \cdot & \cdot \\ \cdot & \mathbf{H}(t) \\ \cdot & \cdot \\ y(t+n+t^*) & \cdot \end{array} \right] , \tag{4.16b}$$

where $\mathbf{H}(t) \in \mathbb{R}^{(t^*+1)\times 2n}$ consists of the last n columns of $\mathbf{G}(t)$, and $\{y(t)\}$ is the actual plant output corresponding to the input $\{u(t)\}$. Because of the plant dynamics relationship (4.7) , we have that

$$\mathbf{G}(t)\, \tilde{\boldsymbol{\Theta}} = 0 \quad , \ \forall\, t \in \mathbb{N}_+ \tag{4.17}$$

Next, because of (4.12),

$$\mathbf{H}(t)^T\, \mathbf{H}(t) \geq \delta\, \mathbf{I} , \ \forall\, t \in \mathbb{N}_+ \tag{4.18}$$

which shows that $\mathbf{H}(t)$ is of maximum column rank. Now, suppose that for some $t_0 \geq t^*+n$, $\boldsymbol{\Theta} = [\,\mathbf{a}^T\ \mathbf{b}^T\,]^T \in \mathbf{A}(t_0)$. Then, for all $t \in \mathbb{N}_+$, $t \leq t_0$, we obtain from the definition of $\mathbf{A}(t_0)$, that

$$| \langle \mathbf{z}(t), \mathbf{a} \rangle - \langle \mathbf{u}(t), \mathbf{b} \rangle | \leq \alpha \ , \tag{4.19}$$

and hence that for all $t \in \mathbb{N}_+$, $t \le t_0$,

$$|\langle \mathbf{u}(t), \mathbf{b}\rangle| \le \alpha + |\langle \mathbf{z}(t), \mathbf{a}\rangle| \ . \tag{4.20}$$

Let $\mathbf{U}(0) \in \mathbb{R}^{(t^*+1)\times n}$ be the submatrix of $\mathbf{H}(0)$ defined by

$$\mathbf{U} \triangleq \begin{bmatrix} \mathbf{u}(0)^T \\ \vdots \\ \mathbf{u}(t^*)^T \end{bmatrix} . \tag{4.21}$$

Then, by (4.20),

$$\|\mathbf{U}\,\mathbf{b}\|_\infty \le \alpha + \max_{t \le t^*} |\langle \mathbf{z}(t), \mathbf{a}\rangle| \ . \tag{4.22}$$

Since $\mathbf{H}(0)$ has maximum column rank, so does the submatrix $\mathbf{U}$. Hence, since $\|\mathbf{a}\|_1 = 1$ by definition of $A(t_0)$, it follows that $A(t_0)$ is bounded, and (i) follows by (iii) of Proposition 4.1.

Next, we proceed to obtain a bound on $\rho(t_0)$, the radius of $A(t_0)$. Let $\mathbf{v} \in A^{\pi}(t_0)$ be arbitrary , and let $\Theta = [\,\mathbf{a}^T\ \mathbf{b}^T\,]^T \in A(t_0)$ be such that $\langle\, \Theta - \mathbf{v}, \mathbf{v} \rangle = 0$. By (4.16a) and (4.19) we have that

$$\|\mathbf{G}(t)\,\Theta\|_\infty \le 2\alpha, \quad \forall\ 0 \le t \le t_0 - k^* - n \ . \tag{4.24}$$

Making use of the relationship between the $\|\cdot\|_2$ and the $\|\cdot\|_\infty$ norms, we now obtain that

$$\|\mathbf{G}(t)\,\Theta\|_2 \le 2\alpha\sqrt{t^*+1}\,, \quad \forall\ \ 0 \le t \le t_0 - t^* - n \tag{4.25}$$

Now, for any $t \in \mathbb{N}_+$, $\mathbf{H}(t)$ has full column rank, and therefore $\mathbf{G}(t)$ can have a null space , $\ker[\mathbf{G}(t)]$, of dimension at most 1 . But $\mathbf{G}(t)\,\tilde{\Theta} = 0$, where $\tilde{\Theta}$ is the true parameter vector. Hence $\ker[\mathbf{G}(t)]$ is spanned by $\tilde{\Theta}$. Furthermore, from (4.18) , $\mathbf{H}(t)$ has has $2n$ singular values greater than $\delta > 0$. By the Courant-Fischer characterization of eigenvalues of Hermitian matrices [11], $\mathbf{G}(t)$ (see (4.16b)) also has $2n$ singular values greater than δ . Let σ_i , $i = 1, 2, \cdots 2n$ be the nonzero singular values of $\mathbf{G}(t)$, with $\sigma_1 \ge\ \sigma_2 \ge\ \cdots\ \ge \sigma_{2n} \ge \delta > 0$. It follows that for all $t \in \mathbb{N}_+$ there exists a singular value decomposition [11]

$$U(t)\Sigma(t)V(t)^T = G(t) \tag{4.26a}$$

where

$$U(t) \in \mathbb{R}^{t^* \times t^*};\ V(t) \in \mathbb{R}^{(2n+1)\times(2n+1)},$$

are orthogonal matrices,

$$\Sigma(t) = \begin{bmatrix} \sigma_1 & 0 & . & 0 & | & 0 \\ 0 & \sigma_2 & . & . & | & . \\ . & . & . & 0 & | & . \\ 0 & . & 0 & \sigma_{2n} & | & 0 \\ - & - & - & - & | & - \\ 0 & . & . & 0 & | & 0 \end{bmatrix} \in \mathbb{R}^{t^* \times (2n+1)}, \tag{4.26b}$$

and such that $V(t)$ is partitioned as follows: $V(t) = [V_1(t) \mid v_{2n+1}]$, with $v_{2n+1} = \tilde{\Theta} / \|\tilde{\Theta}\|_2$.

Now, for any $t \in \mathbb{N}_+$, Θ can be expressed in terms of the orthonormal basis that forms the columns of $V(t)$, i.e.

$$\Theta = V(t)\gamma(t) = [V_1(t) \mid v_{2n+1}] \begin{bmatrix} \gamma_1(t) \\ \gamma_{2n+1} \end{bmatrix} \tag{4.27}$$

for some $\gamma(t) \in \mathbb{R}^{2n+1}$. Since $V_1(t)^T v_{2n+1} = 0$, we must have that $v = V_1(t)\gamma_1(t)$. Now, from (4.25), (4.26) and (4.27) we obtain that for all $t \in \mathbb{N}_+$ with $t \le t_0 - t^* - n$,

$$2\alpha\sqrt{t^*+1} \ge \|G(t)\,\Theta\|_2 = \|U(t)\Sigma(t)V(t)^T V(t)\gamma(t)\|_2 = \|\Sigma(t)\gamma(t)\|_2$$

$$\ge \delta \sqrt{\sum_{i=1}^{2n} \gamma_i(t)^2} = \delta \|V_1(t)\gamma_1(t)\|_2 = \delta \|v\|_2 .$$

because $V_1(t)$ has orthonormal columns. This completes the proof of Lemma 4.1. ∎

>From Lemma 4.1 we see that the persistency of excitation of $\{\varphi(t)\}$ implies that the computation of the plant parameter vector from input - output data has a finite condition number.

Theorem 4.1: Suppose that (i) the sequence $\{\varphi(t)\}$, defined in (4.8), is p.e., (ii) the disturbance sequence $\{d(t)\}$ is a sequence of independent random variables uniformly distributed on $[-\alpha, \alpha]$. Then, $A_\infty = \bigcap_{t=0}^{\infty} A(t) = \{\tilde{\Theta}, -\tilde{\Theta}\}$ with probability 1.

■

Proof: Suppose that there exists a $\Theta \in A_\infty$ such that $\Theta \notin \{\tilde{\Theta}, -\tilde{\Theta}\}$. Then, there exist $\mathbf{w} \in \mathbb{R}^{2n+1}$ and $\eta \in \mathbb{R}$ such that $\mathbf{w} \neq 0$ and

$$\Theta = \mathbf{w} + \eta \mathbf{v}_{2n+1}, \quad \mathbf{v}_{2n+1} = \tilde{\Theta} / \|\tilde{\Theta}\|_2 \text{ and } \langle \mathbf{v}_{2n+1}, \mathbf{w} \rangle = 0 . \tag{4.28}$$

Furthermore, there exists an $\varepsilon > 0$ (without loss of generality, $\varepsilon < \alpha$) such that

$$|[\mathbf{y}(t)^T \mid -\mathbf{u}(t)^T]\mathbf{w}| \geq \varepsilon \text{ infinitely often (i.o.)}. \tag{4.29}$$

To see this, suppose, by way of contradiction, that for any $\varepsilon > 0$, $|[\mathbf{y}(t)^T \mid -\mathbf{u}(t)^T]\mathbf{w}| < \varepsilon$ only finitely often. Then, $\lim_{t\to\infty} |[\mathbf{y}(t)^T \mid -\mathbf{u}(t)^T]\mathbf{w}| = 0$ and therefore (4.16a) implies that $\lim_{t\to\infty} \|\mathbf{G}(t)\mathbf{w}\|_2 = 0$. But $\{\varphi(t)\}$ is p.e. and therefore there exist $t^* \in \mathbf{N}_+$ and $\delta > 0$ such that (4.15) holds. Making use of the singular value decomposition (4.26), we obtain that

$$\begin{aligned} \|\mathbf{G}(t)\mathbf{w}\|_2 &= \|\mathbf{U}(t)\Sigma(t)\mathbf{V}(t)^T\mathbf{w}\|_2 \\ &= \|\Sigma(t)[\mathbf{V}_1(t) \mid \mathbf{v}_{2n+1}]^T\mathbf{w}\|_2 = \left\|\Sigma(t)\begin{bmatrix}\mathbf{V}_1(t)^T\mathbf{w} \\ 0\end{bmatrix}\right\|_2 \\ &\geq \delta\|\mathbf{V}_1(t)^T\mathbf{w}\|_2 = \delta\|\mathbf{w}\|_2 > 0, \end{aligned} \tag{4.30}$$

and therefore $\lim_{t\to\infty} \|\mathbf{G}(t)\mathbf{w}\|_2$ cannot be 0, so that (4.29) must hold.

Now, by the definition of the sets $A(t)$ and A_∞, it follows that

$$|[\mathbf{z}(t)^T \mid -\mathbf{u}(t)^T]\Theta| \leq \alpha, \quad \forall t \in \mathbf{N}_+$$

Substituting (4.28) in the above we obtain

$$-\alpha \le [\mathbf{y}(t)^T \mid -\mathbf{u}(t)^T]\mathbf{w} + \langle \mathbf{d}(t), \mathbf{a} \rangle \le \alpha \ , \quad \forall t \in \mathbb{N}_+,$$

which, together with (4.29) can be used to establish that either

$$\langle \mathbf{d}(t), \mathbf{a} \rangle \ge -\alpha + \varepsilon \tag{4.31a}$$

or

$$\langle \mathbf{d}(t), \mathbf{a} \rangle \le \alpha - \varepsilon \tag{4.31b}$$

or both must hold for infinitely many $t \in \mathbb{N}_+$. Without loss of generality, we may assume that (4.31b) holds for infinitely many $t \in \mathbb{N}_+$. We thus obtain that

$$1 - \mathbf{Prob}\{ A_\infty = \{ \tilde{\theta}, -\tilde{\theta} \} \} \le \mathbf{Prob}\{ \langle \mathbf{d}(t), \mathbf{a} \rangle \le \alpha - \varepsilon \quad \text{i.o.} \} \ . \tag{4.32}$$

Now, for all $t \in \mathbb{N}_+$, if $sgn(a_k) d(t+n-k) > \alpha - \varepsilon$, $k = 0,1,..n$, then,

$$\langle \mathbf{d}(t), \mathbf{a} \rangle = \sum_{k=0}^{n} a_k \, d(t+n-k)$$

$$= \sum_{k=0}^{n} | a_k | \, sgn(a_i) d(t+n-k) > (\alpha - \varepsilon) \sum_{k=0}^{n} | a_k | = \alpha - \varepsilon$$

It follows that, for all $t \in \mathbb{N}_+$,

$$\mathbf{Prob}\{ \langle \mathbf{d}(t), \mathbf{a} \rangle \le \alpha - \varepsilon \} \le 1 - \mathbf{Prob}\{ sgn(a_i) d(t+n-k) > \alpha - \varepsilon \ , i = 0, 1, \ldots n \} \tag{4.33}$$

But $\{ d(t) \}$ is a sequence of independent random variables, uniformly distributed on $[-\alpha, \alpha]$ and therefore for all $t \in \mathbb{N}_+$

$$\mathbf{Prob}\{ \langle \mathbf{d}(t), \mathbf{a} \rangle \le \alpha - \varepsilon \} \le 1 - \prod_{k=0}^{n} \mathbf{Prob}\{ sgn(a_k) d(t+n-k) > \alpha - \varepsilon \} \tag{4.34}$$

$$= 1 - \left(\frac{\varepsilon}{2\alpha} \right)^{n+1} \triangleq p \in (0, 1) \ .$$

If $\langle \mathbf{d}(t), \mathbf{a} \rangle \le \alpha - \varepsilon$ i.o. , then,there exists a subsequence $\{ \mathbf{d}(t_i) \}$ of $\{ \mathbf{d}(t) \}$ such that for all $i \in \mathbb{N}_+$, $t_{i+1} - t_i > n+1$ and $\langle \mathbf{d}(t_i), \mathbf{a} \rangle \le \alpha - \varepsilon$. But

$$\mathbf{Prob}\{ \langle \mathbf{d}(t), \mathbf{a} \rangle \le \alpha - \varepsilon \quad \text{i.o.} \}$$

$$\le \mathbf{Prob}\{ \langle \mathbf{d}(t_i), \mathbf{a} \rangle \le \alpha - \varepsilon \ , \ \forall i \in \mathbb{N}_+ \}$$

$$\leq \mathbf{Prob}\{\langle \mathbf{d}(t_i), \mathbf{a}\rangle \leq \alpha-\varepsilon \ , \ 1 \leq i \leq j\} \ , \ \forall j \in \mathbf{N}_+$$

$$= \prod_{i=0}^{j} \mathbf{Prob}\{\langle \mathbf{d}(t_i), \mathbf{a}\rangle \leq \alpha-\varepsilon\} \leq p^j \ , \ \forall j \in \mathbf{N}_+ \ ,$$

so $\mathbf{Prob}\{\langle \mathbf{d}(t), \mathbf{a}\rangle \leq \alpha-\varepsilon \text{ i.o.}\} = 0$, because $p \in (0,1)$, and this completes the proof.

∎

Since $\{\varphi(t)\}$ cannot be measured, condition (4.15) cannot be checked. We therefore need to develop conditions on the exogenous signals $\{r(t)\}$ and $\{d(t)\}$ which ensure that the sequence $\{\varphi(t)\}$ is p.e. For this purpose, we need the following definitions and useful Lemmas.

Definition 4.5.: A sequence $\{\eta(t)\} \subset \mathbf{R}$ is p.e. *of order n* if $\{\eta(t)\} \subset \mathbb{R}^n$ defined by $\eta(t) = [\eta(t+n-1) \cdots \eta(t)]^T$, $t \in \mathbf{N}_+$ is p.e.

∎

Definition 4.6: A sequence $\{\xi(t)\} \subset \mathbb{R}^n$ has a *spectral line* at $\omega \in [0, 2\pi)$, if there exists a vector $\hat{\xi}(j\omega) \in \mathbf{C}-\{0\}$ such that

$$\lim_{t^* \to \infty} \frac{1}{t^*} \sum_{t=k}^{k+t^*} \xi(t) e^{-j\omega t} = \hat{\xi}(j\omega) \tag{4.35}$$

uniformly in $k \in \mathbf{N}$.

∎

Note that a signal $\{\xi(t)\}$ has a spectral line at $\omega \in [0, 2\pi)$ if it has finite non-zero average energy at the frequency ω .

The continuous time version of the following result can be found in [8].

Lemma 4.2: If a sequence $\{\xi(t)\} \subset \mathbb{R}^N$ has N linearly independent spectral lines $\{\hat{\xi}(j\nu_1) \ldots \hat{\xi}(j\nu_N)\}$, at $\nu_i \in [0, 2\pi)$, $i = 1,2,\ldots,n$, then $\{\xi(t)\}$ is p.e.

The following result is easy to obtain and therefore the proof will be omitted.

Lemma 4.3: If $\{u(t)\} \subset \mathbf{R}$ has N distinct spectral lines $\{\hat{u}(i\nu_1) \cdots \hat{u}(j\nu_N)\}$ at $\nu_i \in [0, 2\pi)$, $i = 1,2,..n$, then $\{u(t)\}$ is p. e. of order N.

The proof of the following result is given in the Appendix.

Lemma 4.4: Consider the state equation

$$\xi(t+1) = \mathbf{A}_c \xi(t) + \mathbf{b}_c u(t), \quad \mathbf{A}_c \in \mathbf{R}^{N \times N}, \quad \mathbf{b}_c \in \mathbf{R}^N \tag{4.36}$$

where the pair $(\mathbf{A}_c, \mathbf{b}_c)$ is completely controllable. If the input $\{u(t)\}$ is p.e. of order N, then, the state $\{\xi(t)\}$ is p.e.

■

An examination of the proof of Lemma 4.1 (in the Appendix) reveals that, for a given input $\{u(t)\}$, the smaller the condition number of the controllability matrix of the system (4.36), the smaller the δ in the definition of persistency of excitation (4.12) will be. Eventually, this observation translates to the conclusion that the closer are the plant polynomial n_p, d_p to having a common factor, the harder it will be to identify the plant parameter $\tilde{\Theta}$.

Theorem 4.2: Consider the feedback system in Figure 1. Suppose that

(i) $\{r(t)\}, \{d(t)\}$ are bounded and that $\{r(t) - d(t)\}$ has $3n+m-1$ spectral lines;

(ii) $C(t,z) \triangleq C_i(z) = \dfrac{n_{c,i}(z)}{d_{c,i}(z)}$ for $t_i \le t \le t_{i+1}$, $i \in \mathbb{N}_+$ where $n_{c,i}(z) = 1 + c_{1,i} z^{-1} + \cdots + c_{m,i} z^{-m}$ and $d_{c,i}(z) = d_{1,i} z^{-1} + \cdots + d_{m,i} z^{-m}$. The compensator update times t_i are defined by the recursion

$$t_0 = 0, \tag{4.37}$$

$$t_{i+1} = t_i + 2n + \min\{ k \in \mathbf{N}_+ \mid \sum_{t=t_i}^{t_i+k} \begin{bmatrix} \mathbf{u}(t+n) \\ \mathbf{u}(t) \end{bmatrix} [\mathbf{u}(t+n)^T \mid \mathbf{u}(t)^T] \ge \gamma I \}$$

for some fixed $\gamma > 0$.

(iii) The sequences $\{[c_{1,i}, \cdots, c_{m,i}]^T\}$ and $\{[d_{1,i}, \cdots, d_{m,i}]^T\}$ are bounded in $\mathbb{R}^n$ with $\{[d_{1,i}, \cdots, d_{m,i}]^T\}$ bounded away from zero. (Note that when the compensator is designed by using (2.7), the above bound conditions on the compensator coefficients can be automatically ensured).

Then, the sequence $\{\varphi(t)\}$ is p.e.

■

Proof:

First, note that if we let

$$\mathbf{A}_c \triangleq \begin{bmatrix} -\frac{\tilde{a}_1}{\tilde{a}_0} & . & . & -\frac{\tilde{a}_n}{\tilde{a}_0} & \frac{\tilde{b}_1}{\tilde{a}_0} & . & . & \frac{\tilde{b}_n}{\tilde{a}_0} \\ 1 & . & . & 0 & 0 & . & . & . \\ 0 & . & . & . & . & . & . & . \\ . & . & 1 & . & 0 & . & . & 0 \\ 0 & . & . & 0 & 1 & . & . & . \\ . & . & . & . & . & . & . & . \\ . & . & . & . & . & . & . & . \\ 0 & . & . & 0 & 0 & . & 1 & 0 \end{bmatrix} \in \mathbb{R}^{2n \times 2n} \tag{4.38}$$

$$\mathbf{b}_c \triangleq \left[0_n \mid 1 \mid 0_{n-1} \right] \in \mathbb{R}^{2n}$$

we have

$$\varphi(t+1) = \mathbf{A}_c\, \varphi(t) + \mathbf{b}_c\, u(t+n) \tag{4.39}$$

and, since $n_p(\tilde{\Theta}, z)$, $d_p(\tilde{\Theta}, z)$ are coprime $(\mathbf{A}_c, \mathbf{b}_c)$ is a completely controllable pair [3]. In view of Lemma 4.4, if $\{u(t)\}$ is p.e. of order n, then $\{\varphi(t)\} \subset \mathbb{R}^{2n}$ is p.e.

Intuitively, if one were to fix the compensator in Fig.1 and assume that all the signals in the loop are bounded, one would see that, because the transfer function from $r-d$ to u has only $n+m-1$ zeros, $\{u(t)\}$ could not have less than $2n$ spectral lines. By Lemma 4.3, it would then follow that $\{u(t)\}$ is p.e. of order $2n$.

We will now present a rigurous argument for the case of the time varying compensator. By Lemma 4.2, condition (i) of Theorem 4.2 implies that the sequence $\{r(t)-d(t)\}$ is p.e. of order $3n+m-1$. Let $\varepsilon(t) \triangleq r(t)-d(t)$ and $\boldsymbol{\varepsilon}(t) \triangleq [\varepsilon(t+3n+m-2) \cdots \varepsilon(t)] \in \mathbb{R}^{3n+m-1}$, $t \in \mathbb{N}_+$. Then, there exist $t^* \in \mathbb{N}_+$, $\delta > 0$, such that forall $\tau \in \mathbb{N}_+$,

$$\sum_{t=\tau}^{\tau+t^*} \boldsymbol{\varepsilon}(t)\boldsymbol{\varepsilon}(t)^T \geq \delta I \ . \tag{4.40}$$

To show that $\{u(t)\}$ is p.e. of order $2n$, it is enough to show that the sequence $\{T_i\}$, defined by $T_i \triangleq t_{i+1}-t_i$, $i \in \mathbb{N}_+$, is bounded . Suppose, by way of contradiction, that $\{T_i\}$ is not bounded. Then, for all $t \in \mathbb{N}_+$ such that $t_i \leq t < t_{i+1}$ and all $i \in K \triangleq \{i \in \mathbb{N}_+ \mid t_{i+1}-t_i > t^* + 3n + m - 1\}$ we have that the transfer function from ε to u is given by

$$H^*_{u,\varepsilon}(\mathbf{x}_i, z) = \frac{d_p(\tilde{\theta}, z)\, n_c(\mathbf{x}_i, z)}{n_c(\mathbf{x}_i, z)\, n_p(\tilde{\theta}, z) + d_c(\mathbf{x}_i, z)\, d_p(\tilde{\theta}, z)} \tag{4.41}$$

$$\triangleq \frac{r_{1,i} z^{-1} + r_{2,i} z^{-2} + \cdots + r_{n+m,i} z^{-(n+m)}}{1 + p_{1,i} z^{-1} + \cdots + p_{n+m,i} z^{-(n+m)}}$$

and therefore, for $t \in \mathbb{N}_+$, $t_i \leq t < t_{i+1}-t^*-(3n+m-1)$, $i \in K$ we have that

$$\sum_{k=0}^{m+n} p_{k,i}\, u(t+m+n-k) = \sum_{k=1}^{m+n} r_{k,i}\, \varepsilon(t+m+n-k) \triangleq v(t) \ . \tag{4.42}$$

Using (4.42) and the notation

$$\mathbf{v}(t) \triangleq [v(t+2n-1)\, v(t+2n-2) \cdots v(t)]^T \in \mathbb{R}^{2n} \tag{4.43a}$$

$$\tilde{\mathbf{u}}(t) \triangleq [u(t+2n-1)\, u(t+2n-2) \cdots u(t)]^T \in \mathbb{R}^{2n} \tag{4.43b}$$

$$\mathbf{V}_{t^*}(t) \triangleq [\mathbf{v}(t+t^*+n+m) \mid \mathbf{v}(t+t^*+m+n-1) \mid \cdots \mid \mathbf{v}(t+m+n)] \in \mathbb{R}^{2n \times (t^*+1)} \tag{4.43c}$$

$$\mathbf{U}_{t^*}(t) \triangleq [\tilde{\mathbf{u}}(t+t^*+m+n) \mid \tilde{\mathbf{u}}(t+t^*+m+n-1) \mid \cdots \mid \tilde{\mathbf{u}}(t)] \in \mathbb{R}^{2n \times (t^*+m+n+1)} \tag{4.43d}$$

$$\mathbf{J}_i \triangleq \begin{bmatrix} r_{1,i} & r_{2,i} & \cdot & r_{m+n-1,i} & r_{n+m,i} & 0 & \cdot & 0 \\ 0 & r_{1,i} & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & 0 \\ 0 & \cdot & 0 & r_{1,i} & r_{2,i} & \cdot & r_{m+n-1,i} & r_{m+n,i} \end{bmatrix} \in \mathbb{R}^{2n \times (3n+m-1)} \ , \tag{4.43e}$$

$$
\mathbf{L}_{t^*,i} \triangleq \begin{bmatrix} 1 & 0 & \cdot & \cdot\;\cdot & 0 \\ p_{1,i} & 1 & 0 & \cdot\;\cdot & \cdot \\ p_{2,i} & p_{1,i} & 1 & \cdot\;\cdot & \cdot \\ \cdot & p_{2,i} & p_{1,i} & \cdot\;\cdot & 0 \\ \cdot & \cdot & p_{2,i} & \cdot\;\cdot & 1 \\ p_{n+m,i} & \cdot & \cdot & \cdot\;\cdot & p_{1,i} \\ 0 & p_{n+m,i} & \cdot & \cdot\;\cdot & p_{2,i} \\ \cdot & 0 & p_{n+m,i} & \cdot\;\cdot & \cdot \\ \cdot & \cdot & 0 & \cdot\;\cdot & \cdot \\ \cdot\;\cdot & \cdot & \cdot & \cdot\;\cdot & \cdot \\ 0 & \cdot & \cdot & \cdot\;0 & p_{n+m,i} \end{bmatrix} \in \mathbb{R}^{(t^*+m+n+1)\times(t^*+1)} \tag{4.43f}
$$

we obtain that for all $t \in \mathbf{N}_+$, $t_i \le t < t_{i+1}-t^*-(3n+m-1)$, $i \in K$

$$
\mathbf{v}(t+n+m) = \mathbf{J}_i\, \varepsilon(t) \ . \tag{4.44}
$$

$$
\mathbf{V}_{t^*}(t) = \mathbf{U}_{t^*}(t)\, \mathbf{L}_{t^*,i}
$$

and

$$
\begin{aligned}
&\lambda_{\min}[\mathbf{J}_i \mathbf{J}_i^T]\, \delta\, \mathbf{I} \qquad\qquad (4.45) \\
&\le \sum_{\tau=t}^{t+t^*} \mathbf{J}_i\, \varepsilon(\tau)\, \varepsilon(\tau)^T \mathbf{J}_i^T \\
&= \sum_{\tau=t+m+n}^{t^*+t+m+n} \mathbf{v}(\tau)\, \mathbf{v}(\tau)^T \\
&= \mathbf{V}_{t^*}(t)\, \mathbf{V}_{t^*}(t)^T \\
&= \mathbf{U}_{t^*,i}(t)\, \mathbf{L}_{t^*,i}\, \mathbf{U}_{t^*,i}(t)^T\, \mathbf{L}_{t^*,i}^T \\
&\le \lambda_{\max}[\mathbf{L}_{t^*,i}\, \mathbf{L}_{t^*,i}]\, \mathbf{U}_{t^*}(t)\, \mathbf{U}_{t^*}(t)^T
\end{aligned}
$$

Since $\mathbf{L}_{t^*,i}$ is full column rank, $\lambda_{\min}[\mathbf{L}_{t^*,i}\, \mathbf{L}_{t^*,i}^T] > 0$, we have that for all $t \in \mathbf{N}_+$, $t_i \le t < t_{i+1}-t^*-(3n+m-1)$, $i \in K$

$$
\sum_{\tau=t}^{t+t^*+m+n} \mathbf{u}(\tau)\, \mathbf{u}(\tau)^T = \mathbf{U}_{t^*}(t)\, \mathbf{U}_{t^*}(t)^T \ \ge\ \frac{\delta\, \lambda_{\min}[\mathbf{J}_i\, \mathbf{J}_i^T]}{\lambda_{\max}[\mathbf{L}_{t^*,i}\, \mathbf{L}_{t^*,i}^T]} \tag{4.46}
$$

By Lemma A.1 (see the Appendix) , there exists a $\beta > 0$ such that for all $i \in \mathbf{N}_+$,

$$\frac{\delta \lambda_{\min}[J_i J_i^T]}{\lambda_{\max}[L_{t^*,i} L_{t^*,i}^T]} \geq \beta .$$

Now, let $l \in N_+$ be such that $l \beta \delta \geq 2\gamma$. Since we assumed that $\{T_i\}$ is unbounded, there exists a k such that $t_{k+1}-t_k > l(t^*+m+n+1)+2n$. It follows from (4.46), that

$$\sum_{t=t_k}^{t_{k+1}-2n} \begin{bmatrix} u(t+n) \\ u(t) \end{bmatrix} [\, u(t+n)^T \mid u(t)^T \,] = \sum_{t=t_k}^{t_{k+1}-2n} \tilde{u}(t)\tilde{u}(t)^T$$

$$\geq \sum_{t=t_k}^{t_k+l(t^*+m+n+1)} \tilde{u}(t)\tilde{u}(t)^T \geq l \beta \gamma \geq 2\gamma ,$$

which contradicts part (ii) of the Theorem's hypothesis. Therefore $\{t_{i+1}-t_i\}$ is a bounded sequence and $\{u(t)\}$ is p.e. of order 2n.

■

Making use of Theorem 3.3 and Theorem 4.2 we obtain the following

Corollary 4.1: Suppose that the conditions (ii) in Theorem 4.1 and (i), (ii) and (iii) in Theorem 4.2 are satisfied. Then the feedback system shown in Fig.1 is uniformly asymptotically stable with probability 1.

■

Since we propose to redesign the control system compensator by semi-infinite optimization using a worst case problem formulation, it is clear from the theorems and discussion in Section 3 that an asymptotically stable, time varying closed loop system can be obtained even if the uncertainty sets $A(t_i)$, used in redesign, do not converge to the doublet $\{-\tilde{\Theta}, \tilde{\Theta}\}$. In addition, by interpreting the results of Theorems 3.2, 3.3 and Lemma 4.1 and Theorem 4.2, we see that this will be the case when the input signal $r(t)$ is rich enough and the disturbance is small enough, as stated below.

Theorem 4.3: Suppose $p(\{\tilde{\Theta}\},\tilde{\mathbf{w}}) < 0$ for a particular choice of weights $\tilde{\mathbf{w}} > 0$, and that the assumptions of Theorem 4.2 are satisfied for $C_i(z) = C(\mathbf{x}_i, z)$, with $\mathbf{x}_i$ a solution of (2.7) for $\mathbf{w} = \tilde{\mathbf{w}}$. Then there exist a $t^* \in \mathbf{N}$ and an $\alpha_0 \in (0,\infty)$ such that if the bound on the disturbance satisfies $\alpha \le \alpha_0$, then (i) $p(A(t),\tilde{\mathbf{w}}) \le 0$ for all $t \ge t^*$ and (ii) the resulting time varying control system, which is obtained by updating the compensator at the times t_i defined in (4.37) using the solutions $\mathbf{x}_i$ of (2.7), is asymptotically stable.

■

If we assume, in addition, the conditions of Theorem 4.1, we need no longer postulate the existence of a bound α_0. the following additional result.

Corollary 4.1: Suppose $p(\tilde{\Theta},\tilde{\mathbf{w}}) < 0$ for a particular choice of weights $\tilde{\mathbf{w}} > 0$, that assumption (ii) of Theorem 4.1 is satisfied and that the assumptions of Theorem 4.2 are satisfied for $C_i(z) = C(\mathbf{x}_i, z)$, with $\mathbf{x}_i$ a solution of (2.7) for $\mathbf{w} = \tilde{\mathbf{w}}$. Then there exist a $t^* \in \mathbb{N}$ such that (i) $p(A(t),\tilde{\mathbf{w}}) \le 0$ for all $t \ge t^*$ and (ii) the resulting time varying control system, which is obtained by updating the compensator at the times t_i defined in (4.37) using the solutions $\mathbf{x}_i$ of (2.7), is asymptotically stable with probability 1.

■

Appendix

Proof of Lemma 4.4

Let $\chi(\lambda) = \lambda^N + \alpha_1\lambda^{N-1} + \cdots + \alpha_N$ be the characteristic polynomial of $\mathbf{A}_c$ and for all $t \in \mathbb{N}_+$, define $\xi(t) \triangleq \mathbf{x}(t+N) + \alpha_1\mathbf{x}(t+N-1) + \cdots + \alpha_N\,\mathbf{x}(t)$. Then, since $\mathbf{x}(t+j) = \mathbf{A}_c^j\,\mathbf{x}(t) + \sum_{i=1}^{j} \mathbf{A}_c^{j-i}\mathbf{b}_c\,\mathbf{u}(t+i-1)$, we have that

$$\xi(t) = \mathbf{A}_c^N\mathbf{x}(t) + \sum_{i=1}^{N} \mathbf{A}_c^{N-i}\mathbf{b}_c\,u(t+i-1) \tag{A.1}$$

$$+ \alpha_1 \mathbf{A}_c^{N-1} \mathbf{x}(t) + \alpha_1 \sum_{i=1}^{N-1} \mathbf{A}_c^{N-1-i} \mathbf{b}_c \, u(t+i-1)$$

$$\vdots$$

$$+ \alpha_{N-1} \mathbf{A}_c \, \mathbf{x}(t) + \alpha_{N-1} \mathbf{b}_c \, u(t)$$

$$+ \alpha_N \mathbf{I} \mathbf{x}(t)$$

$$= [\, \mathbf{b}_c \mid \mathbf{A}_c \mathbf{b}_c + \alpha_1 \mathbf{b}_c \mid \cdots \mid \mathbf{A}_c^{N-1} \mathbf{b}_c + \cdots + \alpha_{N-1} \mathbf{b}_c \,] \begin{bmatrix} u(t+N-1) \\ \vdots \\ u(t) \end{bmatrix}$$

$$= \mathbf{R} \mathbf{T} \mathbf{u}(t) \tag{A.2}$$

where we have used the Cayley-Hamilton Theorem, and the notation

$$\mathbf{R} \triangleq [\, \mathbf{b}_c \mid \mathbf{A}_c \mathbf{b}_c \mid \cdots \mid \mathbf{A}_c^{N-1} \mathbf{b}_c \,] \in \mathbb{R}^{N \times N}$$

for the controllability matrix and

$$\mathbf{T} \triangleq \begin{bmatrix} 1 & \alpha_1 & \alpha_2 & . & . & . & \alpha_{N-1} \\ 0 & 1 & \alpha_1 & . & . & . & . \\ . & 0 & 1 & . & . & . & . \\ . & . & 0. & . & . & . & \alpha_2 \\ . & . & . & . & . & . & \alpha_1 \\ 0 & .0 & 0 & . & . & . & 1 \end{bmatrix} \in \mathbb{R}^{N \times N} \quad \text{and} \quad \mathbf{u}(t) \triangleq \begin{bmatrix} u(t+N-1) \\ u(t+N-2) \\ \vdots \\ u(t) \end{bmatrix} \in \mathbb{R}^N$$

Now, since $\{\mathbf{u}(t)\}$ is p.e. of order N, it follows that there exist $t^* \in \mathbb{N}_+$ and $\delta > 0$ such that

$$\forall\, \tau \in \mathbb{N}_+ , \quad \sum_{t=\tau}^{t=\tau+t^*} \mathbf{u}(t) \mathbf{u}(t)^T \geq \delta \mathbf{I} \; . \tag{A.3}$$

For all $\tau \in \mathbb{N}_+$, let

$$\Xi(\tau) \triangleq [\, \xi(\tau) \mid \xi(\tau+1) \cdots \xi(\tau+t^*) \,] \in \mathbb{R}^{N \times t^*} \tag{A.4a}$$

$$\mathbf{X}(\tau) \triangleq [\, \mathbf{x}(\tau) \mid \mathbf{x}(\tau+1) \cdots \mathbf{x}(\tau+t^*+N) \,] \in \mathbb{R}^{N \times (t^*+N+1)} \tag{A.4b}$$

Then,

$$\Xi(t) = \mathbf{X}(t)\mathbf{M} \tag{A.4c}$$

where

$$\mathbf{M} \triangleq \begin{bmatrix} \alpha_N & 0 & \cdots & 0 \\ \alpha_{N-1} & \alpha_N & \cdots & \cdot \\ \cdot & \alpha_{N-1} & \cdots & \cdot \\ \cdot & \cdot & \cdots & \cdot \\ \cdot & \cdot & \cdots & 0 \\ \alpha_1 & \cdot & \cdots & \alpha_N \\ 1 & \alpha_1 & \cdots & \alpha_{N-1} \\ 0 & 1 & \cdots & \cdot \\ \cdot & 0 & \cdots & \cdot \\ \cdot & \cdot & \cdots & \cdot \\ \cdot & \cdot & \cdots & \alpha_1 \\ 0 & 0 & \cdot\,\cdot\, 0 & 1 \end{bmatrix} \in \mathbb{R}^{(t^*+N+1)\times(t^*+1)}$$

Using, in order (A.4b), (A.4c), (A.4a), (A.2) we obtain that, for all $\tau \in \mathbb{N}_+$, we have :

$$\begin{aligned} \sum_{t=\tau}^{\tau+t^*+N} \mathbf{x}(t)\mathbf{x}(t)^T &= \mathbf{X}(\tau)\mathbf{X}(\tau)^T \\ &\geq \frac{1}{\lambda_{max}[\mathbf{M}\mathbf{M}^T]} \mathbf{X}(\tau)\mathbf{M}\mathbf{M}^T\mathbf{X}(\tau)^T \\ &= \frac{1}{\lambda_{max}[\mathbf{M}\mathbf{M}^T]} \Xi(\tau)\Xi(\tau)^T \\ &= \frac{1}{\lambda_{max}[\mathbf{M}\mathbf{M}^T]} \sum_{t=\tau}^{\tau+t^*} \xi(t)\xi(t)^T \\ &= \frac{1}{\lambda_{max}[\mathbf{M}\mathbf{M}^T]} \mathbf{R}\mathbf{T}\left[\sum_{t=\tau}^{\tau+t^*} \mathbf{u}(t)\mathbf{u}(t)^T\right]\mathbf{T}^T\mathbf{R}^T \\ &\geq \frac{\lambda_{min}[\mathbf{T}\mathbf{T}^T]\lambda_{min}[\mathbf{R}\mathbf{R}^T]}{\lambda_{max}[\mathbf{M}\mathbf{M}^T]} > 0\,, \end{aligned} \tag{A.5}$$

since $(\mathbf{A}_c, \mathbf{b}_c)$ is a completely controllable pair and $\mathbf{T}$ is obviously of full rank . Hence the sequence $\{\mathbf{x}(t)\}$ is p.e. and this completes the proof of Lemma 4.4.

∎

Lemma A.2: Suppose that the assumption (iii) of Theorem 4.2 is satisfied. Then, with $\mathbf{J}_i$, $\mathbf{L}_{t^*,i}$, $i \in \mathbb{N}_+$ as in (4.43e), (4.43f), respectively, we have that

(i) $\inf_{i \in \mathbb{N}_+} \lambda_{\min}[\mathbf{J}_i \mathbf{J}_i^T] > 0$

(ii) $\sup_{i \in \mathbb{N}_+} \lambda_{\max}[\mathbf{L}_{i \bullet,i} \mathbf{L}_{i \bullet,i}^T] < \infty$.

Proof:

The second statement is clear. We shall only show (i) . For all $i \in \mathbb{N}_+$, we have that

$$\mathbf{r}_i = \mathbf{A}_s \mathbf{d}_i$$

where

$$\mathbf{r}_i \triangleq [r_{1,i} | r_{2,i} \cdots r_{m+n,i}]^T \in \mathbb{R}^{m+n}$$

$$\mathbf{d}_i \triangleq [d_{1,i} | d_{2,i} \cdots d_{m,i} | 0 \cdots 0]^T \in \mathbb{R}^{m+n}$$

$$\mathbf{A}_s \triangleq \begin{bmatrix} \tilde{\alpha}_0 & 0 & . & 0 \\ \tilde{\alpha}_1 & \tilde{\alpha}_0 & . & . \\ . & \tilde{\alpha}_1 & . & 0 \\ \tilde{\alpha}_n & . & . & \tilde{\alpha}_0 \\ 0 & \tilde{\alpha}_n & . & \tilde{\alpha}_1 \\ . & . & . & . \\ 0 & . & . & \tilde{\alpha}_n \end{bmatrix} \in \mathbb{R}^{(n+m)\times(n+m)}.$$

But $\mathbf{A}_s$ is of full rank and $\sigma_{\min}(\mathbf{A}_s) > 0$. It follows that

$$\|\mathbf{r}_i\|_2 \geq \sigma_{\min}(\mathbf{A}_s) \|\mathbf{d}_i\|_2$$

and therefore $\{\|\mathbf{r}_i\|_2\}$ is bounded away from 0 . Furthermore, since $\{\|\mathbf{d}_i\|_2\}$ is bounded, it follows that $\{\|r_i\|_2\}$ is bounded . Now, for all $i \in \mathbb{N}_+$, $\mathbf{J}_i$ is of full rank and therefore , for all $i \in \mathbb{N}_+$, $\lambda_i^J \triangleq \lambda_{\min}[\mathbf{J}_i \mathbf{J}_i^T] > 0$. Suppose , by way of contradiction, that there exists a subsequence $\{\lambda_i^J\}_{i \in K}$ of $\{\lambda_i^J\}$ such that $\lim_{i \to \infty,\, i \in K} \lambda_i^J = 0$. Since $\{\mathbf{r}_i\}$ is bounded and bounded away from zero, it follows that there exist an $\hat{\mathbf{r}} \in \mathbb{R}^{m+n} - \{0\}$ and a subsequence $\{\mathbf{r}_i\}_{i \in L}$, $L \subset K$ such that $\lim_{i \to \infty,\, i \in L} \mathbf{r}_i = \hat{\mathbf{r}}$. Then, since $\mathbf{J} \to \lambda_{\min}[\mathbf{J}\mathbf{J}^T]$, $\mathbf{J} \in \mathbb{R}^{2n \times (3n+m+1)}$ is continuous, we obtain that $\lambda_{\min}[\hat{\mathbf{J}}\hat{\mathbf{J}}^T] = 0$ where

$$\mathfrak{J} \triangleq \begin{bmatrix} \hat{r}_1 & \hat{r}_2 & \cdot & \hat{r}_{m+n-1} & \hat{r}_{n+m} & 0 & \cdot & 0 \\ 0 & \hat{r}_1 & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & 0 \\ 0 & \cdot & 0 & \hat{r}_1 & \hat{r}_2 & \cdot & \hat{r}_{m+n-1} & \hat{r}_{m+n} \end{bmatrix} \in \mathbb{R}^{2n \times (3n+m-1)} , \tag{A.6}$$

and is clearly of full row rank, so there exists a neighborhood of 0 that contains only finitely many λ_i^j-s . Hence, there exists a $\beta_1 > 0$ such that $\lambda_{\min}[\mathbf{J}_i \mathbf{J}_i^T] \geq \beta_1 , \forall i \in \mathbf{N}_+$. This completes the proof of (i).

■

Appendix

Proof of Lemma 4.4

Let $\chi(\lambda) = \lambda^N + \alpha_1 \lambda^{N-1} + \cdots + \alpha_N$ be the characteristic polynomial of $\mathbf{A}_c$ and for all $t \in \mathbb{N}_+$, define $\xi(t) \triangleq \mathbf{x}(t+N) + \alpha_1 \mathbf{x}(t+N-1) + \cdots + \alpha_N \mathbf{x}(t)$. Then, since $\mathbf{x}(t+j) = \mathbf{A}_c^j \mathbf{x}(t) + \sum_{i=1}^{j} \mathbf{A}_c^{j-i} \mathbf{b}_c \, \mathbf{u}(t+i-1)$, we have that

$$\xi(t) = \mathbf{A}_c^N + \sum_{i=1}^{N} \mathbf{A}_c^{N-i} \mathbf{b}_c \, u(t+i-1) \tag{A.1}$$

$$+ \alpha_1 \mathbf{A}_c^{N-1} \mathbf{x}(t) + \alpha_1 \sum_{i=1}^{N-1} \mathbf{A}_c^{N-1-i} \mathbf{b}_c \, u(t+i-1)$$

$$\vdots$$

$$+ \alpha_{N-1} \mathbf{A}_c \, \mathbf{x}(t) + \alpha_{N-1} \mathbf{b}_c \, u(t)$$

$$+ \alpha_N \mathbf{I} \mathbf{x}(t)$$

$$= [\, \mathbf{b}_c \mid \mathbf{A}_c \mathbf{b}_c + \alpha_1 \mathbf{b}_c \mid \cdots \mid \mathbf{A}_c^{N-1} \mathbf{b}_c + \cdots + \alpha_{N-1} \mathbf{b}_c \,] \begin{bmatrix} u(t+N-1) \\ \vdots \\ u(t) \end{bmatrix}$$

$$= \mathbf{R} \mathbf{T} \mathbf{u}(t) \tag{A.2}$$

where we have used the Cayley-Hamilton Theorem, and the notation

$$\mathbf{R} \triangleq [\mathbf{b}_c \mid \mathbf{A}_c\mathbf{b}_c \mid \cdots \mid \mathbf{A}_c^{N-1}\mathbf{b}_c] \in \mathbb{R}^{N\times N}$$

for the controllability matrix and

$$\mathbf{T} \triangleq \begin{bmatrix} 1 & \alpha_1 & \alpha_2 & \cdots & \alpha_{N-1} \\ 0 & 1 & \alpha_1 & \cdots & \cdot \\ \cdot & 0 & 1 & \cdots & \cdot \\ \cdot & \cdot & 0. & \cdots & \alpha_2 \\ \cdot & \cdot & \cdot & \cdots & \alpha_1 \\ 0 & 0 & 0 & \cdots & 1 \end{bmatrix} \in \mathbb{R}^{N\times N} \quad \text{and} \quad \mathbf{u}(t) \triangleq \begin{bmatrix} u(t+N-1) \\ u(t+N-2) \\ \vdots \\ u(t) \end{bmatrix} \in \mathbb{R}^N$$

Now, since $\{\mathbf{u}(t)\}$ is p.e. of order N , it follows that there exist $t^* \in \mathbb{N}_+$ and $\delta > 0$ such that

$$\forall\, \tau \in \mathbb{N}_+ , \sum_{t=\tau}^{t=\tau+t^*} \mathbf{u}(t)\mathbf{u}(t)^T \geq \delta \mathbf{I} , \tag{A.3}$$

For all $\tau \in \mathbb{N}_+$, let

$$\Xi(\tau) \triangleq [\,\xi(\tau) \mid \xi(\tau+1) \cdots \xi(\tau+t^*)\,] \in \mathbb{R}^{N\times t^*} \tag{A.4a}$$

$$\mathbf{X}(\tau) \triangleq [\,\mathbf{x}(\tau) \mid \mathbf{x}(\tau+1) \cdots \mathbf{x}(\tau+t^*+N)\,] \in \mathbb{R}^{N\times(t^*+1)} \tag{A.4b}$$

Then,

$$\Xi(t) = \mathbf{X}(t)\mathbf{M} \tag{A.4c}$$

where

$$\mathbf{M} \triangleq \begin{bmatrix} \alpha_N & 0 & \cdots & 0 \\ \alpha_{N-1} & \alpha_N & \cdots & \cdot \\ \cdot & \alpha_{N-1} & \cdots & \cdot \\ \cdot & \cdot & \cdots & \cdot \\ \cdot & \cdot & \cdots & 0 \\ \alpha_1 & \cdot & \cdots & \alpha_N \\ 1 & \alpha_1 & \cdots & \alpha_{N-1} \\ 0 & 1 & \cdots & \cdot \\ \cdot & 0 & \cdots & \cdot \\ \cdot & \cdot & \cdots & \cdot \\ \cdot & \cdot & \cdots & \alpha_1 \\ 0 & 0 & \cdots 0 & 1 \end{bmatrix} \in \mathbb{R}^{(t^*+N+1)\times(t^*+1)}$$

Using, in order (A.4b), (A.4c), (A.4a), (A.2) we obtain that , for all $\tau \in \mathbb{N}_+$, we have :

$$\sum_{t=\tau}^{\tau+\xi^*+N} \mathbf{x}(t)\mathbf{x}(t)^T = \mathbf{X}(\tau)\mathbf{X}(\tau)^T \tag{A.5}$$

$$\geq \frac{1}{\lambda_{max}[\mathbf{M}\mathbf{M}^T]} \mathbf{X}(\tau)\mathbf{M}\mathbf{M}^T\mathbf{X}(\tau)^T$$

$$= \frac{1}{\lambda_{max}[\mathbf{M}\mathbf{M}^T]} \Xi(\tau)\Xi(\tau)^T$$

$$= \frac{1}{\lambda_{max}[\mathbf{M}\mathbf{M}^T]} \sum_{t=\tau}^{\tau+\xi^*} \xi(t)\xi(t)^T$$

$$= \frac{1}{\lambda_{max}[\mathbf{M}\mathbf{M}^T]} \mathbf{R}\mathbf{T}\left[\sum_{t=\tau}^{\tau+\xi^*} \mathbf{u}(t)\mathbf{u}(t)^T\right]\mathbf{T}^T\mathbf{R}^T$$

$$\geq \frac{\lambda_{min}[\mathbf{T}\mathbf{T}^T]\lambda_{min}[\mathbf{R}\mathbf{R}^T]}{\lambda_{max}[\mathbf{M}\mathbf{M}^T]} > 0 ,$$

since $(\mathbf{A}_c, \mathbf{b}_c)$ is a completely controllable pair and $\mathbf{T}$ is obviously of full rank . Hence the sequence $\{\mathbf{x}(t)\}$ is p.e. and this completes the proof of Lemma 4.4.

∎

Lemma A.2: Suppose that the assumption (iii) of Theorem 4.2 is satisfied. Then, with $\mathbf{J}_i$, $\mathbf{L}_{\xi^*,i}$, $i \in \mathbb{N}_+$ as in (4.43e), (4.43f), respectively, we have that

(i) $\inf_{i \in \mathbb{N}_+} \lambda_{min}[\mathbf{J}_i \mathbf{J}_i^T] > 0$

(ii) $\sup_{i \in \mathbb{N}_+} \lambda_{max}[\mathbf{L}_{\xi^*,i} \mathbf{L}_{\xi^*,i}^T] < \infty$.

∎

Proof:

The second statement is clear. We shall only show (i) . For all $i \in \mathbb{N}_+$, we have that

$$\mathbf{r}_i = \mathbf{A}_s \mathbf{d}_i$$

where

$$\mathbf{r}_i \triangleq [r_{1,i} | r_{2,i} \cdots r_{m+n,i}]^T \in \mathbb{R}^{m+n}$$

$$\mathbf{d}_i \triangleq [d_{1,i} | d_{2,i} \cdots d_{m,i} | 0 \cdots 0]^T \in \mathbb{R}^{m+n}$$

$$
\mathbf{A}_s \triangleq \begin{bmatrix} \tilde{\alpha}_0 & 0 & . & 0 \\ \tilde{\alpha}_1 & \tilde{\alpha}_0 & . & . \\ . & \tilde{\alpha}_1 & . & 0 \\ \tilde{\alpha}_n & . & . & \tilde{\alpha}_0 \\ 0 & \tilde{\alpha}_n & . & \tilde{\alpha}_1 \\ . & . & . & . \\ 0 & . & . & \tilde{\alpha}_n \end{bmatrix} \in \mathbb{R}^{(n+m)\times(n+m)}.
$$

But $\mathbf{A}_s$ is of full rank and $\sigma_{\min}(\mathbf{A}_s) > 0$. It follows that

$$
\|\mathbf{r}_i\|_2 \geq \sigma_{\min}(\mathbf{A}_s)\|\mathbf{d}_i\|_2
$$

and therefore $\{\|\mathbf{r}_i\|_2\}$ is bounded away from 0. Furthermore, since $\{\|\mathbf{d}_i\|_2\}$ is bounded, it follows that $\{\|r_i\|_2\}$ is bounded. Now, for all $i \in \mathbb{N}_+$, $\mathbf{J}_i$ is of full rank and therefore, for all $i \in \mathbb{N}_+$, $\lambda_i^J \triangleq \lambda_{\min}[\mathbf{J}_i \mathbf{J}_i^T] > 0$. Suppose, by way of contradiction, that there exists a subsequence $\{\lambda_i^J\}_{i \in K}$ of $\{\lambda_i^J\}$ such that $\lim_{i\to\infty,\, i \in K} \lambda_i^J = 0$. Since $\{\mathbf{r}_i\}$ is bounded and bounded away from zero, it follows that there exist an $\hat{\mathbf{r}} \in \mathbb{R}^{m+n} - \{0\}$ and a subsequence $\{\mathbf{r}_i\}_{i \in L}$, $L \subset K$ such that $\lim_{i\to\infty,\, i \in L} \mathbf{r}_i = \hat{\mathbf{r}}$. Then, since $\mathbf{J} \to \lambda_{\min}[\mathbf{J}\mathbf{J}^T]$, $\mathbf{J} \in \mathbb{R}^{2n\times(3n+m+1)}$ is continuous, we obtain that $\lambda_{\min}[\hat{\mathbf{J}}\hat{\mathbf{J}}^T] = 0$ where

$$
\hat{\mathbf{J}} \triangleq \begin{bmatrix} \hat{r}_1 & \hat{r}_2 & . & \hat{r}_{m+n-1} & \hat{r}_{n+m} & 0 & . & 0 \\ 0 & \hat{r}_1 & . & . & . & . & . & . \\ . & . & . & . & . & . & . & 0 \\ 0 & . & 0 & \hat{r}_1 & \hat{r}_2 & . & \hat{r}_{m+n-1} & \hat{r}_{m+n} \end{bmatrix} \in \mathbb{R}^{2n\times(3n+m-1)}, \tag{A.6}
$$

and is clearly of full row rank, so there exists a neighborhood of 0 that contains only finitely many λ_i^J-s. Hence, there exists a $\beta_1 > 0$ such that $\lambda_{\min}[\mathbf{J}_i \mathbf{J}_i^T] \geq \beta_1$, $\forall\, i \in \mathbb{N}_+$. This completes the proof of (i).

■

5. Conclusion

We have presented two new results. First, we showed that control system design, in a situation requiring tradeoffs, can be very effectively formulated in terms of achievable performance factors that are evaluated by means of semi-infinite optimization algorithms. Second, we developed a new approach to on-line control system tuning, using performance factors, for sequential worst case compensator redesign, together with a new plant uncertainty identification scheme which this approach requires.

In a practical situation, we expect that some measures will be taken to enhance the efficiency of the method described in this paper. Two such measures, in particular, are worth mentioning, in case they should not be obvious to the reader. The first is that it may be possible to have an initial tuning period during which one can apply to the control system a particularly "rich" input since the persistence of excitation constant δ, c.f. (4.12) of the plant input $u(t)$ is proportional to the one of the control system input $r(t)$. Alternatively, the addition of a small "rich" input to a "poor" command input may be be permissible, over an initial period of control system operation.

Another measure which may speed up the uncertainty reduction process is the use of a priori information, such as a knowledge of the location of a pole or zero of the transfer function (2.1). Thus, if z_o is a known pole of the transfer function $P^*(\tilde{\theta},z)$, then $\sum_1^n \tilde{a}_k z_o^{-k} = 0$ must hold. Since this is a linear relationship, it can be added to the definition of the sets $A(t)$ without any effect on the conclusions of the resulting theorems. The effect of adding this relationship, however, is to reduce considerably the size of $A(t)$.

To conclude, we wish to point out the obvious: when new solution tools become available, it becomes possible to formulate and solve design problems that were previously not even considered.

ACKNOWLEDGEMENT

We wish to thank Prof. S. S. Sastry for a number of useful discussions.

This work was supported by the following grants: NSF ECS-8121149, AFOSR 84-0250, ONR N00014-83-K-0602, JSEP F49620-84-C-0057, SRC-82-11-008, and an NSERC (Canada) Postgraduate Scholarship.

5. References

[1] E. Polak, D. Q. Mayne and D. M. Stimler, "Control System Design via Semi-Infinite Optimization: A Review", Proc. IEEE, Vol. 72, No. 12 pp 1777-1794, 1984

[2] E.Polak and D.M. Stimler, "On the Design of Linear Control Systems with Plant Uncertainty via Nondifferentiable Optimization", 9th Triennal IFAC World Congress, Budapest, July 2-6, 1984

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

[4] J.C. Doyle and G. Stein, "Multivariable Feedback Design: Concepts for a Classical/Modern Synthesis", IEEE Trans. Autom. Control, vol. 26, pp. 4-16, 1981

[5] E. Polak, S.E. Salcudean and D. Q. Mayne, "A Sequential Optimal Redesign Procedure for Linear Feedback Systems", U.C.Berkeley ERL Memorandum, to appear

[6] E. Polak and A. N. Payne, "On multicriteria optimization", in Directions in Large Scale Systems, Y. C. Ho and K. S. Mitter, eds., Plenum Press, N.Y., pp 77-94, 1976.

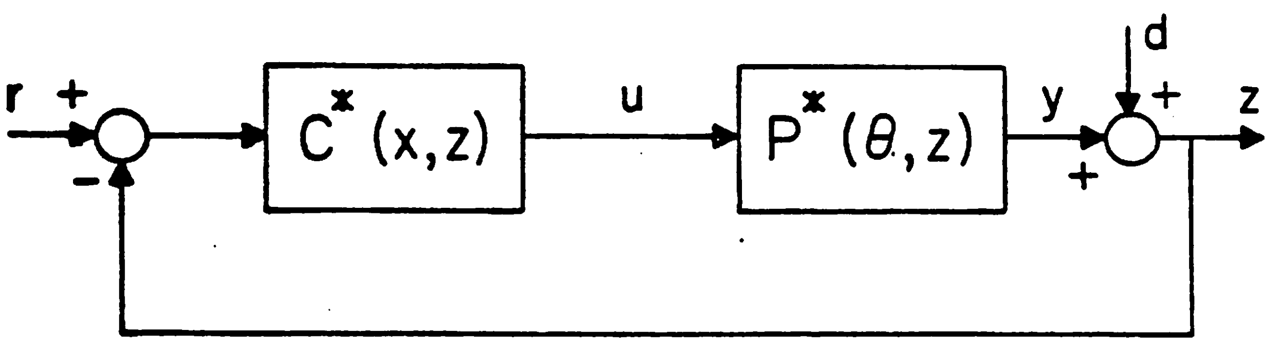
[7] E. Polak, "Notes on the Mathematical Foundations of Nondifferentiable Optimization in Engineering Design", University of California, ERL Memo No. M84/15, Feb. 1984.

[8] S.Boyd and S.Sastry, "Necessary and Sufficient Conditions for Parameter Convergence in Adaptive Control", Proc. of the 1984 American Control Conference, San Diego, June 1984

[9] E. I Jury, *Sampled-Data Control Systems*, John Wiley and Sons, N.Y. 1958

[10] C. Berge, *Topological Spaces*, MacMillan, N. Y., 1963

[11] G. W. Stewart, *Introduction to Matrix Computations*, Academic Press, New York, 1973


r
+
-
$C^*(x,z)$
u
$P^*(\theta,z)$
y
d
+
+
z