FREQUENCY DOMAIN SYNTHESIS OF OPTIMAL INPUTS FOR ADAPTIVE IDENTIFICATION AND CONTROL

by

Li-Chen Fu and Shankar Sastry

Memorandum No. UCB/ERL M87/3

6 February 1987

FREQUENCY DOMAIN SYNTHESIS OF OPTIMAL INPUTS FOR

ADAPTIVE IDENTIFICATION AND CONTROL

by

Li-Chen Fu and Shankar Sastry

Memorandum No. UCB/ERL M87/3

6 February 1987

ELECTRONICS RESEARCH LABORATORY

College of Engineering
University of California, Berkeley
94720

TITLE PAGE

FREQUENCY DOMAIN SYNTHESIS OF OPTIMAL INPUTS FOR

ADAPTIVE IDENTIFICATION AND CONTROL

by

Li-Chen Fu and Shankar Sastry

Memorandum No. UCB/ERL M87/3

6 February 1987

ELECTRONICS RESEARCH LABORATORY

College of Engineering
University of California, Berkeley
94720

Frequency Domain Synthesis of Optimal Inputs for Adaptive Identification and Control

Li-Chen Fu and Shankar Sastry

Electronics Research Laboratory
Department of Electrical Engineering & Computer Science
University of California, Berkeley CA 94720

ABSTRACT

In this paper, we precisely formulate the input design problem of choosing proper inputs for use in SISO Adaptive Identification and Model Reference Adaptive Control algorithms. Characterization of the optimal inputs is given in the frequency domain and is arrived at through the use of *averaging theory*. An expression for what we call the *average information matrix* is derived and its properties are studied. To solve the input design problem, we recast the design problem in the form of an optimization problem which maximizes the smallest eigenvalue of the average information matrix over power constrained signals. A convergent numerical algorithm is provided to obtain the global optimal solution. In the case where the plant has unmodelled dynamics, a careful study of the *robustness* of both Adaptive Identification and Model Reference Adaptive Control algorithms is performed using *averaging theory*. With these results, we derive a bound on the frequency search range required in the design algorithm in terms of the desired performance.

February 6, 1987

Keywords: Averaging, Average Information Matrix, Optimal Input, Sequential Design Algorithm, Convergence Theorem, Adaptive Identification, Model Reference Adaptive Control, Unmodelled Dynamics, Tuned Parameter, Tuned Plant

1. Introduction:

Robustness and *Rate of Convergence* are the two most important factors on the design of adaptive identification and control. To date, there has been a great deal of analysis of robustness properties of adaptive algorithms ever since Rohrs et al [30] showed the lack of robustness margin in several practical applications. While some modifications of adaptive control algorithms have been proposed (Peter & Narendra [29], Kreisselmier & Narendra [20], Sastry [33], Ioannou & Kokotovic [17] and Ioannou & Tsakalis [18]), it has been argued by some that the robustness of adaptive algorithm depends primarily on the persistence of excitation of the controlled system, and, consequently on the choice of the exogenous reference input applied to the system. In Bodson & Sastry [3], the effects of persistent excitation on robustness margin is made precise. A connection between the rate of convergence of the adaptive scheme and the robustness margin is also made. For adaptive identification, Bai & Sastry [2] and Bai et al [12] showed that the parameter converges to a small neighborhood of the *true* parameter of the nominal plant under a condition in the presence of unmodelled dynamics and model mismatch.

The study of parameter convergence rate of adaptive schemes came in the work by Sondhi & Mitra [31] who obtained bounds, dependent on the parameter adaptation gain, of the convergence rate. Later, Fu et al [9], Bodson et al [4] and Kosut et al [21] used averaging techniques to obtain estimates of parameter convergence rate for the nominal system. Such techniques were first introduced by Astrom [1] to explain the instability mechanism in adaptive control arising from unmodelled dynamics, and are the most useful for the analysis of adaptive systems in the frequency domain.

In this paper, we focus on the choice of optimal input signal and its effects on parameter convergence rate in both cases where the plant has and doesn't have unmodelled dynamics. Although the same issue has been discussed by Mareels et al [24], the results are limited. In the stochastic literature, the problem of optimal input design in estimating parameters in linear dynamical systems has been widely investigated (a survey by Mehra [26]). Based on several different optimality criteria such as D-Optimality and A-Optimality, various design algorithms are obtained. Since these designs are unavoidably dependent on the unknown parameters of the system, a Bayesian approach that assumes a prior distribution of the parameters is then used. In all cases, however, the objective of their design is to achieve a more accurate parameter estimate instead of a better parameter convergence rate.

*Research supported by NASA under grant NAG 2-243 and Army Research Office under grant DAAG-29-85-k-0072

The purpose of this paper is both analysis and design. We first analyze the effect of the frequency spectrum of an input to the rate of parameter convergence. For the design purpose, we then find the optimal input (input with the optimal frequency spectrum) by the process of optimizing that effect. Of course, during the course of the optimization, we replace the knowledge of the "true plant" by an estimate. Our contribution is as follows: we propose a *Sequential Design Algorithm* to search for the optimal input signal design which is shown to be the global optimal one. The algorithm is based on the design criterion to maximize the *smallest eigenvalue* of the average information matrix (which will be defined in the sequel) rather than its *determinant* or *trace*. It turns out to be a more practical criterion in the case when the information matrix may be ill-conditioned.

The paper is organized as follows: In section 2, we formulate the input design problem for both Adaptive Identification and Model Reference Adaptive Control in terms of an optimization problem. In section 3, some input rules for design on which the solution of the optimization problem is based are given. In section 4, we propose a numerical design algorithm by which the optimal input is searched sequentially and we give some simulation examples to illustrate the results. In section 5, we discuss the robustness of both Adaptive Identification and Model Reference Adaptive Control schemes. Moreover, we give a bound on the *frequency search range* based on information about the plant uncertainty and the frequency content of the input such that a desirable performance is guaranteed.

2. Input Design Problem

The problem of designing optimal inputs for both Adaptive Identifiers and Model Reference Adaptive Controllers is under investigation.

(I) Adaptive Identifier:

We consider an unknown plant, described by a SISO proper stable transfer function

$$\hat{P}(s) = k_p \frac{n_p(s)}{d_p(s)} \tag{2.1}$$

where $n_p(s)$, $d_p(s)$ are coprime monic polynomials, and $d_p(s)$ is of known degree n.

The adaptive identifier of this plant has a structure shown in Fig. 2.1. The stable filter blocks F_1 and F_2 generate signals $v^{(1)}(t)$ and $v^{(2)}(t)$, which are respectively smoothed derivatives of the input r and the output y_p of the plant. The output of the identifier y_i is obtained through the adaptive gains $c(t)$, $d(t) \in R^n$ and $c_{n+1}(t) \in R$

$$y_i = c^T v^{(1)} + d^T v^{(2)} + c_{n+1} r \tag{2.2}$$

and it may be verified that there exists a unique choice of the adaptive gains, denoted c^*, d^* and c^*_{n+1}, such that the transfer function from the input r to the output y_i is identical to the plant transfer function $\hat{P}(s)$. We define the parameter vector $\theta \in R^{n+1}$

$$\theta^T = (c^T, d^T, c_{n+1}) \tag{2.3}$$

and the signal vector $w \in R^{2n+1}$

$$w^T = (v^{(1)T}, v^{(2)T}, r) \tag{2.4}$$

so that

$$y_i = \theta^T w \tag{2.5}$$

The output of the plant is then given by an equation similar to that of the identifier

$$y_p = \theta^{*T} w \tag{2.6}$$

where θ^* is the vector of "true" parameters corresponding to $\hat{P}(s)$. Defining the parameter error

$$\phi = \theta - \theta^* \tag{2.7}$$

the output error $e_1 = y_i - y_p$ is then given by

$$e_1 = \phi^T w \tag{2.8}$$

It can be shown ([19],[23]) that, with the adaptation law

$$\dot{\phi} = -\Gamma e_1 w \tag{2.9}$$

where $\Gamma \in R^{n \times n}$ is the adaptation gain matrix, the following propositions are true

(i) If $r, \dot{r} \in L_\infty$, then $\lim_{t\to\infty} e_1(t)=0$.

(ii) If, moreover, w is persistently exciting (PE), that is, if there exist constants α_1, α_2 and $\delta>0$ such that

$$\alpha_1 I \le \int_s^{s+\delta} ww^T dt \le \alpha_2 I \qquad \text{for all } s\ge 0 \tag{2.10}$$

then the parameter error also tends to zero, i.e.

$$\lim_{t\to\infty} \phi(t) = 0 \tag{2.11}$$

and the convergence is exponential.

In the proposition (ii) above, we quantify the convergence by giving a preliminary definition.

Definition 2.1: (Exponential Stability, Rate of Convergence)

The equilibrium point $x=0$ of a differential equation is said to be exponentially stable, with rate of convergence α ($\alpha>0$), if

$$\| x(t) \| \le m \| x(t_0) \| e^{-\alpha(t-t_0)} \tag{2.12}$$

for all $t\ge t_0\ge 0$, $x(t_0)\in B_h$ (a closed ball with radius $h>0$) and some $m>1$.

The input design problem for an adaptive identifier is that of selecting an input r from an allowable class of signals (to be specified by the designer) so that the rate of convergence of the parameter error vector ϕ can be optimized. There are various possible solutions to this problem. The solution pursued here is based on a frequency domain approach, applying averaging theory to the update law (2.9), that is to replace Γ by εI where ε is a small positive number. It is shown in [9] that the rate of parameter convergence can be assessed easily by studying the average information matrix $R_w(0)$ defined by

$$R_w(0) = \lim_{T\to\infty} \frac{1}{T} \int_s^{s+T} w(t) w(t)^T dt \qquad s\ge 0 \tag{2.13}$$

when w is persistently exciting. The bound on the rate of convergence is close to the smallest eigenvalue of $R_w(0)$ (a symmetric positive definite matrix) but differs from it by a class k function of ε, $\psi(\varepsilon)$.

The input design problem can therefore be cast in the form of an optimization problem in which an input r is chosen from a class of signals to maximize the smallest eigenvalue of the average information matrix $R_w(0)$. Such a procedure is very reminiscent of the procedure indicated in [14] [16] [25] for the design of input signals in identification. There, however, the objective is to achieve better accuracy of the parameter estimates as opposed to the larger rate of parameter convergence in our case.

(II) Model Reference Adaptive Controller

Next, we examine the optimal input design problem for Model Reference Adaptive Control schemes. We consider the *output error scheme*, developed by Narendra and Valavani [27] and Narendra, Lin and Valavani [28]. Although not discussed here, the *input error scheme*, developed by Bodson and Sastry [5], can be handled similarly.

We consider an SISO plant with transfer function

$$\frac{\hat{y}_p(s)}{\hat{r}(s)} = \hat{P}(s) = k_p \frac{n_p(s)}{d_p(s)} \tag{2.14}$$

where $n_p(s)$ and $d_p(s)$ are monic coprime polynomials of degree m and n respectively and k_p is a scalar. The following are assumed to be known about the plant transfer function:

(A1) The degrees of the polynomials d_p and n_p, namely, n and m, are known.

(A2) The sign of k_p is known (say $k_p > 0$).

(A3) The plant transfer function is assumed to be minimum phase.

The reference model is described by

$$\frac{\hat{y}_m(s)}{\hat{r}(s)} = \hat{M}(s) = k_m \frac{n_m(s)}{d_m(s)} \tag{2.15}$$

where $n_m(s)$ and $d_m(s)$ are monic coprime polynomials of degree m and n respectively (i.e. the same degrees as the corresponding plant polynomials). The reference model is stable, minimum phase, and $k_m > 0$.

The controller structure for the scheme is shown in Fig 2.2. The dynamical compensator blocks F_1 and F_2 (reminiscent of those in the adaptive identifier) are identical single input and n-1 output systems with transfer function $(sI-\Lambda)^{-1}b$ where $\Lambda \in R^{n-1 \times n-1}$, $b \in R^{n-1}$ and Λ is chosen so that its eigenvalues are the zeros of $n_m(s)$. The parameter $c \in R^{n-1}$ in the precompensator block serves to tune the closed loop plant zeros; $d \in R^{n-1}$ and $d_0 \in R$ in the feedback compensator assign the closed loop plant poles. The parameter c_0 adjusts the overall gain of the closed loop plant. Thus, the vector of 2n adjustable parameter denoted θ is

$$\theta^T = (c_0, c^T, d_0, d^T)$$

with the signal vector $w \in R^{2n}$ defined by

$$w^T = (r, v^{(1)T}, y_p, v^{(2)T})$$

The input to the plant is seen to be

$$u = \theta^T w \tag{2.16}$$

and the state equation of the plant loop is given by

$$\begin{bmatrix} \dot{x}_p \\ \dot{v}^{(1)} \\ \dot{v}^{(2)} \end{bmatrix} = \begin{bmatrix} A_p & 0 & 0 \\ 0 & \Lambda & 0 \\ bc_p^T & 0 & \Lambda \end{bmatrix} \begin{bmatrix} x_p \\ v^{(1)} \\ v^{(2)} \end{bmatrix} + \begin{bmatrix} b_p \\ b \\ 0 \end{bmatrix} \theta^T w \tag{2.17}$$

where (A_p, b_p, c_p) is a minimal realization of the plant. It may be verified that there is a unique constant vector $\theta^* \in R^{2n}$ such that, when $\theta = \theta^*$, the transfer function of the plant plus controller equals that of the model, $\hat{M}(s)$.

Now if the relative degree of the plant is one (n-m=1), the model transfer function $\hat{M}(s)$ can be chosen to be strictly positive real, and it can be shown ([27]) that, with the parameter update law

$$\dot{\theta} = -\Gamma e_1 w = -\Gamma(y_p - y_m)w \tag{2.18}$$

where $\Gamma \in R^{2n \times 2n}$ is a positive definite matrix, the following propositions are true:

(i) If the input $r \in L_\infty$, then all signals in the loop, i.e. u, $v^{(1)}$, $v^{(2)}$, y_p and y_m are bounded and

$$\lim_{t \to \infty} e_1(t) = 0 \tag{2.19}$$

(ii) If, moreover, w is persistently exciting (which is similarly defined as in (2.10)), then the parameter error $\phi = \theta - \theta^*$ also tends to zero, i.e.

$$\lim_{t \to \infty} \phi(t) = 0 \tag{2.20}$$

and the convergence is exponential.

The stability proofs of the above propositions heavily rely on the strictly positive realness (SPR) of the model transfer function. In the event that the relative degree n-m is greater than one, $\hat{M}(s)$ can never be chosen to be SPR. Consequently, the following modifications need to be made.

(i) A stable linear filter $L^{-1}(s)$ is found to make the transfer function $L(s)\hat{M}(s)$ SPR.

(ii) When the relative degree n-m is greater than two, augmented output error and over-parametrization (i.e. θ_{2n+1} is used besides $\theta \in R^{2n}$) are used.

It is shown ([6]) that, with a modified update law, propositions (i) and (ii) are still valid (even though the (2n+1)th parameter may not converge in the case when $n - m \geq 3$.).

As a result, the input design problem for this output error control scheme is again that of selecting an input r from a class of signals so as to optimize the rate of parameter convergence.

As in the optimal input design for the adaptive identifier, a frequency domain approach through the application of averaging is adopted as the method to solve the problem. In the following, we will only discuss the case when the relative degree is one. The case in which the relative degree is greater than one can be dealt with similarly.

The state equations for the model loop are given by

$$\begin{bmatrix} \dot{x}_m \\ \dot{v}_m^{(1)} \\ \dot{v}_m^{(2)} \end{bmatrix} = \begin{bmatrix} A_p+b_p d_0^* c_p^T & b_p c^{*T} & b_p d^{*T} \\ b d_0^* c_p^T & \Lambda+bc^{*T} & bd^{*T} \\ bc_p^T & 0 & \Lambda \end{bmatrix} \begin{bmatrix} x_m \\ v_m^{(1)} \\ v_m^{(2)} \end{bmatrix} + \begin{bmatrix} b_p \\ b \\ 0 \end{bmatrix} c_0^* r \tag{2.21}$$

The $(3n-2)\times(3n-2)$ matrix in (2.21) is henceforth referred to as A_m and the $(3n-2)$ vector in (2.21) as b_m. Then subtracting (2.21) from (2.17) with

$$e^T = (x_p^T, v^{(1)T}, v^{(2)T}) - (x_m^T, v_m^{(1)T}, v_m^{(2)T}) \tag{2.22}$$

we have that

$$\dot{e} = A_m e + b_m \phi^T w \tag{2.23}$$

and

$$e_1 = (c_p^T, 0, 0)e := c_m^T e \tag{2.24}$$

Note from (2.21) that $c_0^* c_m^T (sI - A_m)^{-1} b_m$ is equal to the model transfer function $\hat{M}(s)$ and that $c_0^* = \frac{k_m}{k_p}$ is the ratio of the high frequency gains. Now the update law (2.18) becomes

$$\dot{\theta} = \dot{\phi} = -\Gamma w c_m^T e \tag{2.25}$$

To apply averaging, we consider slow adaptation, i.e. $\Gamma = \varepsilon I$ resulting in

$$\dot{e} = A_m e + b_m w^T \phi \tag{2.26}$$

$$\dot{\phi} = -\varepsilon w c_m^T e \tag{2.27}$$

Recall that w is not exogenously specified, rather it depends on e, i.e.

$$w = w_m + Qe \tag{2.28}$$

where w_m is an exogenously defined 3n-2 dimensional vector obtained from r and given by

$$w_m = (r, v_m^{(1)T}, y_m, v_m^{(2)T})$$

and Q is the constant $2n\times(3n-2)$ matrix

$$Q = \begin{bmatrix} 0 & 0 & 0 \\ 0 & I & 0 \\ c_p^T & 0 & 0 \\ 0 & 0 & I \end{bmatrix}$$

Using these, eq. (2.26), (2.27) are rewritten as

$$\dot{e} = A_m e + b_m w_m^T \phi + b_m e^T Q^T \phi \tag{2.29}$$

$$\dot{\phi} = -\varepsilon w_m c_m^T e - \varepsilon Q e c_m^T e \tag{2.30}$$

With the exception of the last terms (quadratic in e and ϕ), eq. (2.29), (2.30) are linear time varying equations describing the linearized adaptive control system, around the equilibrium point e=0 and ϕ=0.

Recall that exponential parameter convergence can be obtained provided that w is persistently exciting. Referring to the definition 2.1, we see this is the case by taking the rate of parameter convergence as that of tail parameter convergence; in other words, behavior of small e and ϕ is our domain of interest. Consequently, we apply the averaging to the *linearized* version of (2.29), (2.30), i.e.

$$\dot{e} = A_m e + b_m w_m^T \phi \tag{2.31}$$

$$\dot{\phi} = -\varepsilon w_m c_m^T e \tag{2.32}$$

The rate of parameter convergence of the linearized adaptive system can be easily assessed by investigating its averaged system

$$\dot{\phi}_{av} = -\varepsilon R_{w_m w_{mf}}(0) \phi_{av} \tag{2.33}$$

where

$$R_{w_m w_{mf}}(0) := \lim_{T\to\infty} \frac{1}{T} \int_s^{s+T} w_m(t) \frac{1}{c_0^*} \hat{M}(w_m)(t) dt \qquad s \geq 0 \tag{2.34}$$

(Note $\frac{1}{c_0^*}\hat{M}(w_m)(t)$ is a vector obtained by filtering $w_m(t)$ through a transfer function $\frac{1}{c_0^*}\hat{M}(s)$)

It is shown in [9] that the bound on the rate of convergence is close to the smallest real part of eigenvalues of $R_{w_m w_{mf}}(0)$ (a positive definite *non-symmetric matrix*) but differs from it by a class k function of ε, $\psi(\varepsilon)$. The resulting optimization problem is to choose an input r (subject to certain constraints) so as to maximize the smallest real part of eigenvalues of the general (non-symmetric) matrix $R_{w_m w_{mf}}(0)$. Although this is not well formed, the next lemma ([8]) remedies the situation.

Lemma 2.1:

Given any constant square real matrix A such that

$$A = H_1 + iH_2 \tag{2.35}$$

where H_1 and H_2 are Hermitian matrices, then the following are true:

$$\lambda_{\min}(H_1) \leq \mathrm{Re}\lambda(A) \leq \lambda_{\max}(H_1) \tag{2.36}$$

$$\lambda_{\min}(H_2) \leq \mathrm{Im}\lambda(A) \leq \lambda_{\max}(H_2) \tag{2.37}$$

where $\lambda(\cdot)$ stands for an eigenvalue of some matrix in its argument. In the case where A is a real matrix, H_1 and iH_2 are simply symmetric and skew-symmetric matrices respectively.

In light of Lemma 2.1, the transformation from input design problem to a problem of maximizing the smallest eigenvalue of the symmetric part of $R_{w_m w_{mf}}(0)$, denoted as $S_{ymm}[R_{w_m w_{mf}}(0)]$, by choosing r subject to some constraints is thus established.

3. Input Design Bases:

The input design problem is formulated in terms of an optimization of the smallest eigenvalue of a positive symmetric matrix (i.e. $R_w(0)$ for the adaptive identifier and $S_{ymm}[R_{w_m w_{mf}}(0)]$ for the Model Reference Adaptive Controller) over a class of input signals. In this section, we make the problem more tractable by choosing the class of input signals to be power-constrained; by which we roughly mean that the average power of a signal $i(t) \in R$, defined as

$$\lim_{T\to\infty} \frac{1}{T} \int_{s}^{s+T} i^2(t)dt \tag{3.1}$$

with limit existing uniformly in $s \geq 0$, can be no greater than a fixed amount. In the following, more detailed definitions are introduced to facilitate the later development of the input design algorithm.

Definition 3.1: (Stationary, Autocovariance, Power Spectral Measure)

A signal $u(\cdot) : R_+ \to R^n$ is said to be stationary if, for each τ, the following limit exists uniformly in s:

$$R_u(\tau) = \lim_{T\to\infty} \frac{1}{T} \int_{s}^{s+T} u(t+\tau)u^T(t)dt \tag{3.2}$$

in which instance, the limit $R_u(\tau)$ is called the *autocovariance* of u. Also, $R_u(\tau)$ can be written as the inverse Fourier Transform of a positive power spectral measure $S_u(d\omega)$

$$R_u(\tau) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{j\omega\tau} S_u(d\omega) \tag{3.3}$$

As indicated by Definition 3.1, the average power of a stationary vector signal u may thus be expressed in terms of its power spectral measure $S_u(d\omega)$, i.e.

$$R_u(0) = \frac{1}{2\pi} \int_{-\infty}^{\infty} S_u(d\omega) \tag{3.4}$$

In general, a stationary signal may have nonzero power spectral measure over continuous and/or discrete sets in the frequency spectrum. The next definition explores this in more detail.

Definition 3.2: (Frequency Support)

The frequency support of a scalar stationary signal u with power spectral measure $S_u(d\omega)$ is defined as

$$\mathrm{Supp}(u) = \left\{ \omega \mid \omega \in R \text{ , for all } \varepsilon > 0 \text{ , } \int_{\omega-\varepsilon}^{\omega+\varepsilon} S_u(d\omega) > 0 \right\} \quad , \tag{3.5}$$

Let $F_u(\omega)$ be defined as

$$F_u(\omega) = \int_{-\infty}^{\infty} S_u(d\omega') \tag{3.6}$$

then $F_u(d\omega)$ is a spectral distribution function which is monotonically increasing and continuous from the right. If $F_u(d\omega)$ is absolutely continuous, then the frequency support Supp(u) defines a continuous spectrum, which denotes the smallest closet set outside which the power spectral measure S_u vanishes. On the other hand, if $F_u(d\omega)$ is a staircase function with n jumps, then Supp(u), which is exactly n points in the frequency spectrum, defines the discrete spectrum.

In practice, stationary signals frequently encountered are bandlimited with bandwidth $\Omega=[-\omega_o,\omega_o]$, $\omega_o>0$. This leads to the following definition.

Definition 3.3: (Normalized Input Design (NID) Set $N(\Omega)$)

A normalized input design is defined by the spectral distribution function $F_u(\omega)$ which satisfies

$$\frac{1}{2\pi} F_u(\infty) = 1 \tag{3.7}$$

The normalized input design set $N(\Omega)$ is then defined to be a set of normalized input designs with frequency support only contained in the frequency band $\Omega=[-\omega_o,\omega_o]$, i.e.

$$F(-\omega_o^-) = 0 \ , \qquad F(\omega_o) = 1 \tag{3.8}$$

Note that $F(\omega)$ can be identified with a positive measure, i.e.

$$\int_\Omega S(d\omega) = \int_\Omega dF(\omega) \tag{3.9}$$

which gives a more concise expression of $N(\Omega)$

$$N(\Omega) = \left\{ F \mid F : \text{positive measure} \ , \ \frac{1}{2\pi}\int_\Omega dF(\omega) = 1 \right\} \tag{3.10}$$

In the sequel, we will use $N_D(\Omega)$ to denote a subset of $N(\Omega)$, including all the NID's with only discrete spectrum contained in Ω.

Definition 3.4: (Normalized Average Information Matrix (NAIM))

A matrix G is said to be a normalized average information matrix if there exists a proper stable column transfer function $H(\cdot):C\rightarrow C^m$, a scalar function $a(\cdot):C\rightarrow R_+$ and $F\in N(\Omega)$ such that

$$G = \frac{1}{2\pi}\int_\Omega a(j\omega)H(j\omega)H^*(j\omega)dF(\omega) \tag{3.11}$$

In the special case where $F(\omega)$ results from a single frequency sinusoidal input with frequency ω_0, the corresponding NAIM G will be called the *point-input information matrix* (PIIM) and denoted $G(\omega_0)$. Moreover, all NAIM's resulting from the column transfer function $H(s)$ and all possible F's in $N(\Omega)$ form a class of matrices $M_H(\Omega)$ called NAIM set, i.e.

$$M_H(\Omega) = \left\{ G \mid G = \frac{1}{2\pi}\int_{\Omega} a(j\omega)H(j\omega)H^{*}(j\omega)dF(\omega) \ , \ F \in N(\Omega) \right\} \tag{3.12}$$

Definition 3.5: (Minimum Eigenvalue of $G \in M_H(\Omega)$)

A function $\lambda_G(\cdot):N(\Omega) \rightarrow R$ is defined to be the minimum eigenvalue of a NAIM G resulting from some NID $F \in N(\Omega)$.

Remarks:

(1) The NID set $N(\Omega)$ is a convex set due to the fact that

$$(1-\alpha)F_1 + \alpha F_2 \in N(\Omega) \tag{3.13}$$

for all $F_1, F_2 \in N(\Omega)$ and $\alpha \in [0,1]$.

(2) All NAIM's are symmetric and at least positive semi-definite, therefore, $\lambda_G(F)$ are nonnegative for all $F \in N(\Omega)$.

(3) In the stochastic literature, the average information matrix that we used here is often referred to as *Fisher information matrix* which is related to the error covariance matrix. Due to the similarity between the average information matrix and the Fisher information matrix, in the following, we will state some lemmas with proofs omitted (they can be obtained as in [16] [25]).

Lemma 3.1:

The NAIM set $M_H(\Omega)$ is the closed convex hull of all PIIM's corresponding to the same transfer function H, i.e.

$$M_H(\Omega) = Co\left\{ G(\omega) \mid \omega \in \Omega \right\} \tag{3.14}$$

Lemma 3.2:

For any $F_1 \in N(\Omega)$ with corresponding $G(F_1) \in M_H(\Omega)$, there always exists a $F_2 \in N_D(\Omega)$ containing no more than m(m+1)/2+1 distinct frequency elements (m(m+1)+2 spectral lines) such that

$$G(\mathrm{F}_1) = G(\mathrm{F}_2) \tag{3.15}$$

(Note that $G(\mathrm{F})$ is denoted to emphasize the dependence of G on F.)

Lemma 3.3:

The optimal normalized input $\mathrm{F}^* = \mathrm{argmax}\ \{\ \lambda_G(\mathrm{F}) \mid \mathrm{F} \in N(\Omega)\ \}$ exists, and contains no more than m(m+1)/2 distinct frequency elements (i.e. one less than that predicted by Lemma 3.2).

Remark:

One can infer from Lemma 3.3 that while designing optimal inputs for maximizing the smallest eigenvalue of the average information matrix, one can confine the search to sinusoidal inputs with a finite number of frequencies.

4. Sequential Design Algorithm and Its Application to Adaptive Identifier and Model Reference Adaptive Controller

In this section, we first derive some basic results on the smallest eigenvalue of $G(F)$, namely, $\lambda_G(F)$, using perturbation theory. The numerical algorithm for input design given later is based on these results. In the end, we illustrate the results by showing simulation examples.

Theorem 4.1: (Equivalence Theorem)

Consider some $F^* \in N(\Omega)$. Let $\lambda_G(F^*)$ be the smallest eigenvalue of $G(F^*)$ and v_i , i=1,...,r be the orthonormal eigenvectors associated with it. Then the following three statements are equivalent.

$$(a) \qquad F^* = \text{argmax}\left\{ \lambda_G(F) \mid F \in N(\Omega) \right\} \tag{4.1}$$

$$(b) \quad \text{for all } F^o \in N(\Omega)\text{ , with } \quad F^\alpha = (1-\alpha)F^* + \alpha F^o\text{ , } \quad \alpha \in [0,1] \tag{4.2}$$

$$\frac{\partial}{\partial \alpha}[\lambda_G(F^\alpha)]|_{\alpha=0} \le 0 \tag{4.3}$$

$$(c) \qquad \lambda_G(F^*) \ge \underline{\sigma}_{max} \tag{4.4}$$

where

$$\underline{\sigma}_{max} = \max\left\{ \lambda(P^T G(F)P) \mid F \in N(\Omega) \right\} \tag{4.5}$$

and

$$P = [v_1, \cdots, v_r] \tag{4.6}$$

Proof of Theorem 4.1:

The way we proceed in the proof is to show (a) (b) are equivalent and then (b) (c) are equivalent.

(i) First of all, note that from (4.2)

$$G(F^\alpha) = (1-\alpha)G(F^*) + \alpha G(F^o) \tag{4.7}$$

and that, by perturbation theory, the smallest eigenvalue satisfies

$$\lambda_G(F^\alpha) = (1-\alpha)\lambda_G(F^*) + \alpha\underline{\sigma} + o(\alpha) \tag{4.8}$$

when α is small, where $\underline{\sigma}$ is defined by

$$\underline{\sigma} = \lambda(P^T G(F^o)P) \tag{4.9}$$

and

$$P = [v_1, \cdots, v_r] \tag{4.10}$$

It then follows that (a) implies (b) trivially. To show that (b) implies (a), we use a contradiction.

Suppose (b) is true but that there exists a $\hat{F} \neq F^*$ such that

$$\underline{\lambda}_G(\hat{F}) > \underline{\lambda}_G(F^*) \tag{4.11}$$

Define F^α as

$$F^\alpha = (1-\alpha)F^* + \alpha\hat{F} \qquad \alpha \in [0,1] \tag{4.12}$$

Then

$$G(F^\alpha) = (1-\alpha)G(F^*) + \alpha G(\hat{F}) \tag{4.13}$$

and the smallest eigenvalue satisfies

$$\underline{\lambda}_G(F^\alpha) = (1-\alpha)\underline{\lambda}_G(F^*) + \alpha\underline{\sigma} + o(\alpha) \tag{4.14}$$

when α is small and $\underline{\sigma}$ is defined by

$$\underline{\sigma} = \underline{\lambda}(P^T G(\hat{F}) P) \tag{4.15}$$

where P is defined as in (4.10).

Since, by definition, v_i, $i=1, \cdots, r$ are orthonormal vectors, one can easily show that

$$\underline{\sigma} \geq \underline{\lambda}_G(\hat{F}) \tag{4.16}$$

Further, with eq. (4.14), one can establish the following

$$\frac{\partial}{\partial\alpha}[\underline{\lambda}_G(F^\alpha)]|_{\alpha=0} = \underline{\sigma} - \underline{\lambda}_G(F^*) \tag{4.17}$$

which along with (4.11) and (4.16) gives a contradiction. Hence, the implication is valid.

(ii)

(c) => (b). By hypothesis and definition of $\underline{\sigma}_{max}$, we have

$$\underline{\lambda}_G(F^*) \geq \underline{\sigma}(F^o) \qquad \text{for all } F^o \in N(\Omega) \tag{4.18}$$

where

$$\underline{\sigma}(F^o) = \underline{\lambda}(P^T G(F^o) P) \tag{4.19}$$

With definition of F^α in (4.2), (4.18) then implies that

$$\frac{\partial}{\partial\alpha}[\underline{\lambda}_G(F^\alpha)]|_{\alpha=0} = \underline{\sigma} - \underline{\lambda}_G(F^*) \leq 0 \tag{4.20}$$

$(b) \Rightarrow (c)$. This is more obvious to see since if $\underline{\sigma}_{max} > \underline{\lambda}_G(F^*)$, then there exists $\bar{F} \in N(\Omega)$ and F^α defined by

$$F^\alpha = (1-\alpha)F^* + \alpha\bar{F} \qquad \alpha \in [0,1] \tag{4.21}$$

such that

$$\frac{\partial}{\partial\alpha}[\underline{\lambda}_G(F^\alpha)]|_{\alpha=0} = \underline{\sigma}(\bar{F}) - \underline{\lambda}_G(F^*) > 0 \tag{4.22}$$

Consequently, (b) and (c) are equivalent.

Q.E.D.

Remark:

In the Equivalence Theorem, one should note that finding $\underline{\sigma}_{max}$ is less complex than finding $\underline{\lambda}_{max} := \{ \underline{\lambda}_G(F) \mid F \in N(\Omega) \}$ in the nontrivial case simply because $P^T G(F) P$ is of dimension $r \times r$ and $r < m$. In fact, the most common and the simplest case is when P consists of single vector where $P^T G(F) P$ becomes a scalar. Thus, by Lemma 3.3, $\underline{\sigma}_{max}$ can be easily calculated using a one-line search optimization routine, i.e.

$$\underline{\sigma}_{max} = \max_{\omega \in \Omega} P^T G(\omega) P \tag{4.23}$$

Except in very simple cases, the computation of optimal inputs has to be done numerically. We propose the following algorithm and prove that it converges to the global optimum. To start with, we introduce some notation that will be used in the sequel.

Notation:

(i) $N_D^k(\Omega)$ is a subset of $N_D(\Omega)$ with each member containing no more than k sinusoidal components.

(ii) $P_i = [v_{i,1}, \cdots, v_{i,r_i}]$ consists of orthonormal eigenvectors of $G(F^i)$ associated with the smallest eigenvalue $\underline{\lambda}_G(F^i)$.

(iii) $\underline{\sigma}^i_{max} = \max \{ \underline{\lambda}(P_i^T G(F) P_i) \mid F \in N_D^{k_i}(\Omega) \}$ where $k_i = r_i(r_i+1)/2$.

Numerical Algorithm:

Data: $F^o \in N_D^{k_o}(\Omega)$ is a feasible initial design.

Step1: Set i = 0.

Step2: Compute $\underline{\lambda}_G(F^i)$ and find $\underline{\sigma}^i_{max}$.

Step3: If $\underline{\sigma}^i_{max} \le \underline{\lambda}_G(F^i)$, then stop and F^i is the optimal input design; else go to step 4.

Step4: Update the input design F^i by

$$F^{i+1} = (1-\alpha_i)F^i + \alpha_i \hat{F}^i \qquad \alpha_i \in [0,1] \tag{4.24}$$

where $\hat{F}^i \in N_D^{k_i}(\Omega)$ is such that

$$\sigma^i_{max} = \lambda_G(P_i^T G(\hat{F}^i) P_i) \tag{4.25}$$

Step5: i = i + 1 and go to step 2.

Remark:

In step 2 of the numerical algorithm, the procedure of finding σ^i_{max} is exactly the same as finding $\lambda_G(F^*)$, i.e. to go through step 2 to step 5 with some feasible initial design $F^{i_o} \in N_D^{k_i}(\Omega)$.

Theorem 4.2: (Convergence Theorem)

In the sequential design algorithm, if the sequence $\{ \alpha_i \}$ is chosen such that

$$(a) \qquad \lim_{i \to \infty} \alpha_i = 0 \qquad \sum_{i=1}^{\infty} \alpha_i = \infty \qquad \alpha_i \in (0,1) \tag{4.26}$$

or

$$(b) \qquad \alpha_i = \text{argmax} \{ \ \lambda_G[(1-\alpha)F^i + \alpha \hat{F}^i] \ \mid \alpha \in [0,1] \ \} \tag{4.27}$$

then either the numerical algorithm terminates in finite steps or

$$\lambda_G(F^i) \to \lambda_G(F^*) \qquad \text{as } i \to \infty \tag{4.28}$$

where F^* is an optimal input design as defined in (4.1).

Remark: The optimal input design as defined in (4.1) is not unique. The input design sequence $\{F^i\}$ generated in the sequential design algorithm will converge to one of the the optimal input designs in the sense of (4.28).

Proof of Theorem 4.2:

Let's assume the algorithm does not stop in finite steps.

(a) Equivalent to showing that (4.28) is true, we show that

$$\frac{\partial}{\partial \alpha} [\lambda_G(F^\alpha_{(i)})]|_{\alpha=0} \le 0 \qquad \text{as } i \to \infty \tag{4.29}$$

where

$$F^\alpha_{(i)} = (1-\alpha)F^i + \alpha \hat{F}^i \tag{4.30}$$

Assume the contrary, i.e.

$$\frac{\partial}{\partial\alpha}[\lambda_G(F^{\alpha}_{(i)})]|_{\alpha=0} = \Delta > 0 \qquad \text{for all } i\geq 0 \tag{4.31}$$

This implies

$$\lim_{i\to\infty} [\ \lambda_G(F^i) - \lambda_G(F^0)\ ] = (\ \sum_{i=1}^{\infty} \alpha_i\ )\ \delta(\Delta) \tag{4.32}$$

where $\delta(\Delta)>0$, which contradicts that $\{\ \lambda_G(F^i)\ \}$ is a bounded sequence.

(b) By the Equivalence theorem, we know that if F^i is not the optimal input design, then

$$\frac{\partial}{\partial\alpha}[\lambda_G(F^{\alpha}_{(i)})]|_{\alpha=0} > 0 \tag{4.33}$$

which yields $\{\ \lambda_G(F^i)\ \}$ as a monotonically increasing sequence bounded above. Hence, the sequence converges to a limit, say, $\lambda_G(\bar{F})$. We now show that

$$\lambda_G(\bar{F}) = \lambda_G(F^*)$$

where F^* is assumed to be an optimal input design.

Assume a contradiction, i.e. $\lambda_G(\bar{F})\neq\lambda_G(F^*)$. Again, by Equivalence Theorem, the gradient

$$\frac{\partial}{\partial\alpha}[\lambda_G(F^{\alpha})]|_{\alpha=0} = \Delta > 0 \tag{4.34}$$

where F^{α} is defined as

$$F^{\alpha} = (1-\alpha)\bar{F}+\alpha F \tag{4.35}$$

for some $F\in N(\Omega)$. This, in turn, implies

$$\lim_{i\to\infty} [\ \lambda_G(F^i) - \lambda_G(F^{i-1})\ ] = \delta(\Delta) > 0 \tag{4.36}$$

which contradicts that the sequence converges. In consequence,

$$\lim_{i\to\infty} \lambda_G(F^i) = \lambda_G(F^*) \tag{4.37}$$

Q.E.D.

Remark:

In effect, by numerically constructing the design F^*, we obtain the design F_s which can be made arbitrarily close to F^* but in general is usually distinct from it (we can take any large but finite number of iterations). Since the design F_s may have undesirably large point spectrum, its approximation is usually considered. Such an approximation, discussed in [14], yields the rounded-off design and will be denoted by F^*_{rnd} subsequently.

To apply the sequential design algorithm to the Adaptive Identifier and Model Reference Adaptive Controller, we initially obtain the spectral representation of $R_w(0)$ and $S_{ymm}[R_{w_m w_{mf}}(0)]$ by use of eq. (2.13) and (2.34). It may be verified that, as a result of applying a normalized input design F(ω), the following are true.

$$R_w(0) = \frac{1}{2\pi}\int n(j\omega)n^*(j\omega)dF(\omega) \tag{4.38}$$

and

$$S_{ymm}[R_{w_m w_{mf}}(0)] = \frac{1}{2\pi}\int \frac{\mathrm{Re}M^*(j\omega)}{c_0^*}\bar{n}(j\omega)\bar{n}^*(j\omega)dF(\omega) \tag{4.39}$$

where $n(s)$ stands for the column transfer function from r to w in the Adaptive Identifier and $\bar{n}(s)$ stands for the column transfer function from r to w_m in the Model Reference Adaptive Controller. By definition 3.4, if F is taken to be an NID in $N(\Omega)$, then it is easy to see that $R_w(0)$ and $S_{ymm}[R_{w_m w_{mf}}(0)]$ are NAIM's in $M_n(\omega)$ and $M_{\bar{n}}(\omega)$ respectively. Application of the numerical algorithm then readily yields the optimal designs for both cases.

To illustrate the preceding results, we show examples in the adaptive identifier and controller respectively, in which instances the plants and the model are the same as in [9]. As for the purpose of design, we compute the PIIM by using an estimate of the unknown parameters.

Example 1: Consider the identification of the plant

$$\hat{P}(s) = \frac{2s+2}{s+3} \tag{4.40}$$

The filter is chosen to be $\det(sI-\Lambda)=(s+5)$. The true values of c_1, d_1, c_2 are -1.6, 0.4, and 2.0. Denote the parameter error as:

$$\phi_1 = c_1-c_1^* \qquad \phi_2 = d_1-d_1^* \qquad \phi_3 = c_2-c_2^* \tag{4.41}$$

To calculate the PIIM $G(\omega)$ required in the numerical algorithm, we use an estimate of the plant transfer function, namely, $\hat{P}_i(s)$

$$\hat{P}_i(s) = \frac{3(s+2)}{(s+5)}$$

and the corresponding initial guess of the parameters are

$$c_1^i = -1.2 \qquad d_1^i = 0.0 \qquad c_2^i = 3.0 \tag{4.42}$$

Thus, based on this estimate, $G(\omega)$ becomes

$$G(\omega) = \begin{bmatrix} \frac{25}{25+\omega^2} & \frac{75(15+\omega^2)}{(25+\omega^2)^2} & \frac{25}{25+\omega^2} \\ \frac{75(15+\omega^2)}{(25+\omega^2)^2} & \frac{225(9+\omega^2)}{(25+\omega^2)^2} & \frac{15(75+7\omega^2)}{(25+\omega^2)^2} \\ \frac{25}{25+\omega^2} & \frac{15(75+7\omega^2)}{(25+\omega^2)^2} & 1 \end{bmatrix} \tag{4.43}$$

Since the number of unknown parameters is 3, parameter convergence will occur when the support of F(ω) consists of greater than or equal to 3 points. Hence, we choose the initial input design as

$$\frac{1}{2\pi}\,F^0 = \frac{1}{2}\delta(\omega)+\frac{1}{4}\delta(\omega-2)+\frac{1}{4}\delta(\omega+2) \tag{4.44}$$

and the frequency search range Ω=[−10,10]. After applying sequential design algorithm, we obtain F^*_{rnd} (rounded-off design)

$$\frac{1}{2\pi}\,F^*_{rnd} = 0.445\delta(\omega)+0.192\{\delta(\omega-2)+\delta(\omega+2)\}+0.0203\{\delta(\omega-3.52)+\delta(\omega+3.52)\}$$

$$+\;0.00702\{\delta(\omega-3.80)+\delta(\omega+3.80)\}+0.00442\{\delta(\omega-4.29)+\delta(\omega+4.29)\}$$

$$+\;0.0539\{\delta(\omega-4.43)+\delta(\omega+4.43)\}+0.10\{\delta(\omega-10)+\delta(\omega+10)\} \tag{4.45}$$

In Fig. 4.1, we show the spectral distribution of F^*_{rnd}, while in Fig. 4.2 (a) (b) (c) and Fig. 4.3 (a) (b) (c), we illustrate the time trajectories of parameter errors ϕ_1, ϕ_2, and the output error $e=y_t-y_p$ for the input designs F^0 and F^*_{rnd} respectively.

Example 2: Consider the adaptive control process of a first order plant

$$\hat{P}(s) = \frac{2}{(s+1)} \tag{4.46}$$

By adjusting the feedforward gain k_r and feedback gain k_y, the closed loop transfer function is made to match the model transfer function

$$M(s) = \frac{3}{(s+3)} \tag{4.47}$$

in which case, the true values of k_r and k_y are 1.5 and -1. Define $\phi_1=k_r-k_r^*$ and $\phi_2=k_y-k_y^*$. The PIIM is then computed to be

$$G(\omega) = \begin{bmatrix} \frac{9}{(9+\omega^2)} & \frac{81}{(9+\omega^2)^2} \\ \frac{81}{(9+\omega^2)^2} & \frac{81}{(9+\omega^2)^2} \end{bmatrix} \tag{4.48}$$

To guarantee the persistency of excitation, we choose the initial input design F^0 to be

$$\frac{1}{2\pi}\,F^0 = \frac{1}{2}\delta(\omega-1.5)+\frac{1}{2}\delta(\omega+1.5) \tag{4.49}$$

Application of sequential design algorithm yields the rounded-off optimal input design F^*_{rnd} as

$$\frac{1}{2\pi}\,F^*_{rnd} = \frac{1}{2}\delta(\omega-2.46)+\frac{1}{2}\delta(\omega+2.46) \tag{4.50}$$

Figure 4.4 (a) (b) (c) and Figure 4.5 (a) (b) (c) display the time trajectories of the parameter errors ϕ_1, ϕ_2 and the output error $e_1 = y_p - y_m$ for the input designs F^0 and F^*_{rnd} respectively.

5. Robustness Discussion: Presence of Unmodelled Dynamics

In section 4, we apply the numerical design algorithm to both Adaptive Identifiers and Model Reference Adaptive Controllers where the plant has no unmodelled dynamics, in which case the frequency search range Ω may be made as large as possible. In fact, however, the adaptive identification and the Model Reference Adaptive Control (MRAC) are usually undertaken in the case where the plant is contaminated by unmodelled dynamics. As a consequence, the choice of the frequency search range Ω becomes a relatively important factor for consideration in the context of input design.

In this section, first, we analyze the robustness of both adaptive identification and MRAC; second, we explicitly derive a bound on the frequency search range based on prior information about the plant (and the model). To start with, consider the following definition:

Definition : (Crosscovariance, Cross-Power Spectral Measure)

Given two stationary signals $u(\cdot) : R_+ \rightarrow R^n$, $v(\cdot) : R_+ \rightarrow R$, $R_{uv}(\tau)$ is called the *crosscovariance* of u and v if

$$R_{uv}(\tau) = \lim_{T\to\infty} \frac{1}{T} \int_{s}^{s+T} u(t+\tau)v(t)dt \tag{5.1}$$

in which case $R_{uv}(\tau)$ can be written as the inverse Fourier Transform of a cross-power spectral measure $S_{uv}(d\omega)$

$$R_{uv}(\tau) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{j\omega\tau} S_{uv}(d\omega) \tag{5.2}$$

As indicated by the definition, the average cross-power of two stationary signals u and v may thus be expressed in terms of its cross-power spectral measure $S_{uv}(d\omega)$, i.e.

$$R_{uv}(0) = \frac{1}{2\pi} \int_{-\infty}^{\infty} S_{uv}(d\omega) \tag{5.3}$$

Next, we will study the robustness of the Adaptive Identifier and the Model Reference Adaptive Controller respectively.

(I) Adaptive Identifier

We consider a finite order plant consisting of the same nominal plant as in section 2 along with some stable additive unmodelled dynamics, i.e.

$$\hat{P}(s) = k_p \frac{n_p(s)}{d_p(s)} = \hat{P}^*(s) + \Delta\hat{P}(s) \tag{5.4}$$

satisfying the following assumptions:

(D1) $|\hat{P}^*(j\omega)| \le L_i$ for some known positive constant L_i and for all $\omega \in R$.

(D2) There exists a known nondecreasing function $l_i(\omega): R \to R_+$ such that $l_i(\omega) \to 0$ as $\omega \to 0$ and

$$|\Delta\hat{P}(j\omega)| \le l_i(\omega) \le l_{i\infty} \qquad \text{for all } \omega \in R \tag{5.5}$$

(D3) The input $r(t)$ is stationary and $r(t),\ \dot{r}(t) \in L_\infty$.

Recall that the signal vector $w \in R^{2n+1}$ is defined to be

$$w^T = (v^{(1)T}, v^{(2)T}, r) \tag{5.6}$$

which can be reproduced, differing only by stable initial condition terms, by filtering the input $r(t)$ through a single-input multi-output linear system with transfer function $n(s)$

$$n^T(s) = \left[\frac{1}{\hat{\Lambda}(s)}, \cdots, \frac{s^{n-1}}{\hat{\Lambda}(s)}, \frac{\hat{P}(s)}{\hat{\Lambda}(s)}, \cdots, \frac{s^{n-1}\hat{P}(s)}{\hat{\Lambda}(s)}, 1\right]$$

$$= \frac{1}{\hat{\Lambda}(s)d_p(s)} \left[d_p(s), \cdots, s^{n-1}d_p(s), n_p(s), \cdots, s^{n-1}n_p(s), \hat{\Lambda}(s)d_p(s)\right] \tag{5.7}$$

where $\hat{\Lambda}(s) = \det(sI - \Lambda)$ is a Hurwitz polynominal with $\partial(\hat{\Lambda}(s)) = n$, and (Λ, b) is the controllable canonical realization of the compensator block as shown in Fig. 2.1.

If the plant does not have unmodelled dynamics, it can be shown ([7]) that $w(t)$ is PE provided the input $r(t)$ contains 2n+1 spectral lines. In other words, the matrix Z_{2n+1} defined by

$$Z_{2n+1} := \left[n(j\omega_1), n(j\omega_2), \cdots, n(j\omega_{2n}), n(j\omega_{2n+1})\right] \tag{5.8}$$

is nonsingular whenever $\omega_i \ne \omega_j$, $i \ne j$, $i,j = 1, 2, \ldots, 2n+1$. On the other hand, in the case when the plant has unmodelled dynamics, it is shown in [12] that $w(t)$ is almost always PE when the input $r(t)$ contains 2n+1 spectral lines or, equivalently, the matrix Z_{2n+1} is nonsingular for almost all $(\omega_1, \omega_2, \cdots, \omega_{2n+1})^T \in R^{2n+1}$. Based on this result, we give the following proposition.

Proposition 5.1:

Given the transfer function $n(s)$ as in (5.7), define the matrix $Z_\gamma \in C^{(2n+1)\times\gamma}$ as

$$Z_\gamma := \left[n(j\omega_1), n(j\omega_2), \cdots, n(j\omega_{\gamma-1}), n(j\omega_\gamma)\right] \tag{5.9}$$

For almost all $(\omega_1, \omega_2, \cdots, \omega_\gamma)^T \in R^\gamma$, $\gamma \le 2n+1$, the matrix Z_γ is of full column rank, i.e. $\rho(Z_\gamma) = \gamma$. Equivalently, the set $\{ (\omega_1, \omega_2, \cdots, \omega_\gamma)^T \in R^\gamma \mid \rho(Z_\gamma) < \gamma \}$ is of zero measure.

Proof: See Appendix.

In fact, a real stationary signal contains symmetric spectral lines, e.g. $j\omega$ and $-j\omega$ when $\omega \neq 0$. Hence, $r(t)$, containing 2n+1 spectral lines, should consist of n sinusoids and a DC signal which gives a set of frequencies $(0,\omega_1,\cdots,\omega_n,-\omega_1,\cdots,-\omega_n)$. To practicalize the results from [12] and the previous proposition, we introduce a technical assumption which, we believe, will be true in general.

Assumption:

(D4) Given $l = 1,\cdots,n$, there exists a set of frequencies $(0,\omega_1,\cdots,\omega_l,-\omega_1,\cdots,-\omega_l)$ such that no nonzero $\xi_l \in R^{2l+1}$

$$\xi_l = \left[\mu_0 , \mu_1 , \cdots , \mu_l , \tau_0 , \tau_1 , \cdots \tau_{l-1} \right]^T$$

will satisfy

$$\hat{\mu}_l(j\omega_i)\, d_p(j\omega_i) + \hat{\tau}_l(j\omega_i)\, n_p(j\omega_i) = 0 \qquad i = -l,\cdots,-1,0,1,\cdots l \tag{5.10}$$

where

$$\hat{\mu}_l(s) = \mu_0 + \mu_1 s + \cdots + \mu_l s^l \qquad \text{and} \qquad \hat{\tau}_l(s) = \tau_0 + \tau_1 s + \cdots + \tau_{l-1} s^{l-1}$$

Remark:

(1) This assumption has been implicitly taken for the work of synthesizing transfer functions using frequency response data, e.g. [32],[34].

(2) When $l = n$, the assumption (D4) implies that the matrix Z_{2n+1} defined by

$$Z_{2n+1} := \left[n(0), n(j\omega_1), n(-j\omega_1), \cdots n(j\omega_n), n(-j\omega_n) \right] \tag{5.11}$$

is nonsingular. In other words, there exist n sinusoids $\sin(\omega_i t)$, i = 1,...,n, such that the signal vector $w(t)$ will be PE if the input is of the form

$$r(t) = a_0 + \sum_{i=1}^{n} a_i \sin(\omega_i t + b_i) \qquad a_i, b_i \in R \tag{5.12}$$

Based on this assumption and the previous results, we give the following proposition which deals with the case of sinusoidal inputs.

Proposition 5.2: (Almost-All Richness of Sinusoidal Input)

Given the transfer function $n(s)$ as in (5.7) satisfying the assumption (D4). Define the matrix $Z_{2l+1} \in C^{(2n+1)\times(2l+1)}$ by

$$Z_{2l+1} := \left[n(0), n(j\omega_1), n(-j\omega_1), \cdots, n(j\omega_l), n(-j\omega_l) \right] \tag{5.13}$$

Then

(i) for almost all $(\omega_1,\omega_2,\cdots,\omega_l)^T \in R^l$, $l \geq n$, the matrix Z_{2l+1} is of full row rank, i.e. $\rho(Z_{2l+1})=2n+1$. Equivalently, the set $\{ (\omega_1,\omega_2,\cdots,\omega_l)^T \in R^l \mid \rho(Z_{2l+1})<2n+1 \}$ is of zero measure.

(ii) for almost all $(\omega_1,\omega_2,\cdots,\omega_l)^T \in R^l$, $l \leq n$, the matrix Z_{2l+1} is of full column rank, i.e. $\rho(Z_{2l+1})=2l+1$. Equivalently, the set $\{ (\omega_1,\omega_2,\cdots,\omega_l)^T \in R^l \mid \rho(Z_{2l+1})<2l+1 \}$ is of zero measure.

Proof: See Appendix.

Corollary 5.3:

Given the transfer function $n(s)$ as in (5.7), for almost all $(\omega_1,\cdots,\omega_l)^T \in R^l$, $l \leq n$, there exists a $\theta_0 \in R^{2n+1}$ such that

$$\hat{P}(j\omega_i) = \theta_0^T n(j\omega_i) \qquad i=-l,\cdots,0,\cdots,l \qquad \omega_{-i}=-\omega_i \text{ and } \omega_0=0 \tag{5.14}$$

and is given by

$$\theta_0 = Z_{2l+1}[Z^*_{2l+1} Z_{2l+1}]^{-1} g_{2l+1} \tag{5.15}$$

where g_{2l+1} is defined to be

$$g^*_{2l+1} := \left[\hat{P}(0), \hat{P}(j\omega_1), \hat{P}(-j\omega_1), \cdots \hat{P}(j\omega_l), \hat{P}(-j\omega_l) \right] \tag{5.16}$$

Proof: See Appendix.

Consequently, a conclusion can be readily drawn from the Corollary 5.3 about the Adaptive Identifier that, for almost all inputs $r(t)$ containing l sinusoids where $l \leq n$ and a DC signal, there always exists an $\theta_0 \in R^{2n+1}$ such that the error signal $e^*(t) := y_p(t) - \theta_0^T w(t)$ converges to zero exponentially in time t. Therefore, the parameter update law defined by (2.9) can be rewritten in a form which helps to analyze the stability, i.e.

$$\dot{\theta} = -\Gamma(y_i - y_p)w = -\Gamma w w^T(\theta - \theta_0) + \Gamma(y_p - \theta_0^T w)w \tag{5.17}$$

or

$$\dot{\phi} = -\Gamma w w^T \phi + \Gamma e^* w \tag{5.18}$$

where $\phi := \theta - \theta_0$. Since the last term on the RHS of (5.18) is an exponentially decaying term, it can be shown that the asymptotic behavior of ϕ in (5.18) is no different from the one below

$$\dot{\phi} = -\Gamma w w^T \phi \tag{5.19}$$

By such fact we can conclude that $\phi^T w \to 0$ as $t \to \infty$ and ,in turn, that $y_i - y_p = \phi^T w + e^* \to 0$ as $t \to \infty$. In addition, the *partial convergence theorem* ([6]) implies that

$$\lim_{t \to \infty} R_w(0)\phi(t) = 0 \tag{5.20}$$

where $R_w(0)$ is as defined in (2.13).

Remark: In particular, when the input $r(t)$ contains exactly n sinusoids and a DC signal, then with almost all n sinusoids, $w(t)$ is PE and the parameter $\theta_0 \in R^{2n+1}$ such that $y_p \to \theta_0^T w$ as $t \to \infty$ is uniquely defined. Moreover, the parameter $\theta(t)$ then converges to θ_0 exponentially; in which case an identified transfer function associated with the parameter θ_0 (which will be called *tuned parameter* as defined in [12]) can be obtained. Though the true plant is stable, the identified plant transfer function could be unstable due to the inappropriate location of the frequencies of the sinusoidal input.

On the other hand, when the input $r(t)$ contains l sinusoids and a DC signal where $l > n$, then for almost all l sinusoids, $w(t)$ is PE but there may not be a solution $\theta \in R^{2n+1}$ to the following set of equations

$$\hat{P}(j\omega_i) = \theta^T n(j\omega_i) \qquad i = -l, \cdots, 0, \cdots, l \qquad \omega_{-i} = -\omega_i \text{ and } \omega_0 = 0 \tag{5.21}$$

Instead, there will be a minimum error (least square error) solution θ_0 corresponding to that family of equations (5.21) and is given by

$$\theta_0 = [Z_{2l+1} Z_{2l+1}^*]^{-1} Z_{2l+1} g_{2l+1} \tag{5.22}$$

After some algebra, the expression for θ_0 in (5.18) becomes

$$\theta_0 = \left[\sum_{i=-l}^{l} n(j\omega_i) n^T(-j\omega_i) \right]^{-1} \left[\sum_{i=-l}^{l} n(j\omega_i) \hat{P}(-j\omega_i) \right] \tag{5.23}$$

The θ_0 so obtained will match the *tuned parameter* θ_T as defined in [12]

$$\begin{aligned} \theta_T &= R_w(0)^{-1} R_{y_p w}(0) \\ &= \left[\sum_{i=-l}^{l} n(j\omega_i) n^T(-j\omega_i) \frac{a_i^2}{t_i} \right]^{-1} \left[\sum_{i=-l}^{l} n(j\omega_i) \hat{P}(-j\omega_i) \frac{a_i^2}{t_i} \right] \qquad a_{-i} = a_i \end{aligned} \tag{5.24}$$

where

$$t_i = \begin{cases} 4 & i \neq 0 \\ 2 & i = 0 \end{cases} \tag{5.25}$$

when the weights of all the spectral lines, $\frac{a_i^2}{t_i}$, i = $-l,\ldots,0,\ldots,l$ are taken to be 1. It is also shown in [12] that the parameter $\theta(t)$ in the parameter update equation

$$\dot{\theta} = -\varepsilon w w^T \theta + \varepsilon y_p w \tag{5.26}$$

is a bounded function of time $t \geq 0$ and

$$\lim_{t \to \infty} \| \theta(t) - \theta_T \| \leq \overline{\psi}(\varepsilon) \tag{5.27}$$

where $\overline{\psi}(\varepsilon)$ is a class k function of ε,

Remarks:

(1) To compare θ_T defined in (5.24) with θ_0, note that

$$\theta_T = [Z_{2l+1} W Z_{2l+1}^*]^{-1} Z_{2l+1} W g_{2l+1} \tag{5.28}$$

where W is a diagonal weighting matrix

$$W = \text{diag}\left(\frac{a_1^2}{t_1}, \frac{a_{-1}^2}{t_{-1}}, \ldots, \frac{a_l^2}{t_l}, \frac{a_{-l}^2}{t_{-l}}\right) \tag{5.29}$$

(2) θ_T may be considered to be a natural tuned parameter due to two facts: (i) the steady-state parameter $\theta(t)$ stays close to θ_T and is different from it only by $\overline{\psi}(\varepsilon)$ when the parameter update is slow as in (5.26) (ii) the steady-state plant output y_p stays close to the identifier output y_i since that

$$| y_i - y_p | \leq | y_p - w^T \theta_T | + \| \theta_T - \theta(t) \| \, \|w\| \tag{5.30}$$

where θ_T is the weighted minimum error solution to (5.21).

Next, we derive a bound on the frequency search range Ω used in the sequential design algorithm in terms of the amount of deviation of θ_T from the nominal parameter θ^* (corresponding to the nominal plant with the transfer function $\hat{P}^*(s)$).

Denote by w^* and y_{p^*} the signal vector and the output corresponding to the nominal plant and rewrite (5.26) by

$$\begin{aligned} \dot{\theta} &= -\varepsilon (w^* + \Delta w)(w^* + \Delta w)^T \theta + \varepsilon (y_{p^*} + \Delta y_p)(w^* + \Delta w) \\ &= -\varepsilon (w^* w^{*T} + \Delta B)\theta + \varepsilon (y_{p^*} w^* + \Delta C) \end{aligned} \tag{5.31}$$

where $\Delta B = w^* \Delta w^T + \Delta w w^{*T} + \Delta w \Delta w^T$ and $\Delta C = y_{p^*} \Delta w + \Delta y_p w^* + \Delta y_p \Delta w$. In analogy to [12], the averaged system of (5.31) can be written as

$$\dot{\theta}_{av} = -\varepsilon R_w(0) \theta_{av} + \varepsilon R_{y_p w}(0)$$

$$= -\varepsilon(R_{w^*}(0)+\Delta R_w(0))\theta_{av}+\varepsilon(R_{y_p w^*}(0)+\Delta R_{y_p w}(0)) \tag{5.32}$$

where

$$R_{w^*}(0) = \lim_{T\to\infty}\frac{1}{T}\int_0^T w^* w^{*T} dt \qquad \Delta R_w(0) = \lim_{T\to\infty}\frac{1}{T}\int_0^T \Delta B dt \tag{5.33}$$

and

$$R_{y_p w^*}(0) = \lim_{T\to\infty}\frac{1}{T}\int_0^T y_p w^* dt \qquad \Delta R_{y_p w}(0) = \lim_{T\to\infty}\frac{1}{T}\int_0^T \Delta C dt \tag{5.34}$$

With the fact that

$$\theta^* = R_{w^*}(0)^{-1}R_{y_p w^*}(0) \tag{5.35}$$

and, from (5.24) and (5.32), that

$$\theta_T = [R_{w^*}(0)+\Delta R_w(0)]^{-1}[R_{y_p w^*}(0)+\Delta R_{y_p w}(0)] \tag{5.36}$$

the following proposition gives the bound on the frequency search range Ω in terms of the bound on $\|\theta_T - \theta^*\|$.

Proposition 5.4: (Bound on Ω in adaptive identification)

Given an Adaptive Identifier satisfying assumptions (D1)-(D4). Suppose the input $r(t)$ is power constrained, containing a DC signal and finite number, say $s \ge n$, of sinusoids with frequencies lying inside the compact set $\Omega = [-\varpi, \varpi]$ in the frequency spectrum, i.e.

$$\frac{1}{2\pi}\int_\Omega dF_r(\omega) = \sum_{i=-s}^{s}\frac{a_i^2}{l_i} = F_\infty \tag{5.37}$$

Let $\beta_1 \le \|\theta^*\| \le \beta_2$ for some $\beta_1, \beta_2 > 0$ and $\|(j\omega I - \Lambda)^{-1}b\| \le K_\Lambda$, $\omega \in R$, for some $K_\Lambda > 0$. If ϖ satisfies

$$0 < \frac{\alpha_1(\varpi)\,\beta_2 + \alpha_2(\varpi)}{\beta_1\,(\,\underline{\lambda}[R_{w^*}(0)] - \alpha_1(\varpi)\,)} \le \eta \tag{5.38}$$

for a small number $\eta > 0$ where

$$\alpha_1(\omega) = \left\{2K_\Lambda\sqrt{1+(1+L_i^2)K_\Lambda^2} + K_\Lambda^2 l_{i\infty}\right\} l_i(\omega)F_\infty \tag{5.39}$$

$$\alpha_2(\omega) = \left\{\sqrt{1+(1+L_i^2)K_\Lambda^2} + K_\Lambda(l_{i\infty}+L_i)\right\} l_i(\omega)F_\infty \tag{5.40}$$

then

$$\|\theta_T - \theta^*\| \le \eta\|\theta^*\| \tag{5.41}$$

Proof: See Appendix.

Remarks:

(1) Despite that the bound on Ω is not explicit, by the nature of the functions α_1 and α_2, ϖ can be obtained numerically.

(2) The derivation of ϖ and, in turn, $\Omega=[-\varpi,\varpi]$ in (5.38) requires the knowledge of $\underline{\lambda}[R_{w^*}(0)]$, which can be replaced by an estimate similar to the one used in the numerical design algorithm.

(3) Normally, the optimal input design generated by the numerical design algorithm consists of more than n sinusoids; in which case the parameter $\theta(t)$ won't converge to θ_T but rather oscillates around it. Hence, the preferable input design should be a two-phase design. The phase I design is simply the optimal input design itself such that $\theta(t)$ converges to the neighborhood of θ_T fast and stays within it. The phase II design is to add one DC signal and reduce the number of sinusoids in the phase I design to exactly n such that $\theta(t)$ settles down on a new θ_T. However, the determination of the time when this phase II design should be initiated and the way that these n sinusoids are to be chosen in the phase II design requires some experience on the part of the designer.

Example 1: Consider the first example in section 4 with the plant being changed into

$$\hat{P}(s) = \frac{2(s+1)}{(s+3)}\frac{30}{(s+30)} \tag{5.42}$$

The corresponding additive unmodelled dynamics is

$$\Delta\hat{P}(s) = \frac{2(s+1)}{(s+3)}\frac{-s}{(s+30)} \tag{5.43}$$

Recall that the nominal parameter is $c_1^*=-1.6$, $d_1^*=0.4$, and $c_2^*=2.0$. Suppose we use the optimal input design obtained in (4.45) as the phase I design, the tuned parameter θ_T can be calculated, using the definition $\theta_T=R_w(0)^{-1}R_{y_pw}(0)$, to be $c_1=-1.531$, $d_1=0.819$, $c_2=1.650$, and the tuned plant transfer function is thus to be

$$\hat{P}_{\bar{\theta}}(s) = \frac{(1.650s+0.593)}{(s+0.905)} \tag{5.44}$$

Now, we start the phase II design at t=20 sec such that the initial design F^0 in the first example in section 4 is used instead of the original design F^*_{nd}. Again, the tuned parameter θ_T and the tuned plant transfer function are found to be $c_1=-1.443$, $d_1=0.516$, $c_2=1.767$ and

$$\hat{P}_{\bar{\theta}}(s) = \frac{(1.767s+1.613)}{(s+2.422)} \tag{5.45}$$

Fig. 5.1 (a) (b) (c) show the time trajectories of the parameter errors ϕ_1, ϕ_2, and the output error $e=y_t-y_p$. After roughly t=40 sec, the adjustable parameter θ settles down on the tuned parameter value. In Fig. 5.2 (a) (b) (c), we show the time trajectories for the parameter error and the output error using only the design F^0. Apparently, the adjustable parameter θ does not converge to the tuned value even when t=160 sec. To illustrate the closeness among the true plant, the nominal plant, and the tuned plant, we draw the Nyquist plots of the transfer functions of these plants and show them in Fig. 5.3.

(II) Model Reference Adaptive Controller

In the case of MRAC, we consider a finite order plant consisting of the same nominal plant as in section 2 (with relative degree one) and some stable multiplicative unmodelled dynamics, i.e.

$$\hat{P}(s) = k_p\frac{n_p(s)}{d_p(s)} = \hat{P}^*(s)(1+U(s)) \tag{5.46}$$

, and the same reference model as in section 2

$$\hat{M}(s) = k_m\frac{n_m(s)}{d_m(s)} \tag{5.47}$$

satisfying the following assumptions:

(E1) $|\hat{M}(j\omega)| \le L_c$ for some known positive constant L_c and for all $\omega \in R$.

(E2) There exist some known nondecreasing functions $l_{c1}(\omega)$ and $l_{c2}(\omega): R \to R_+$ such that $l_{c1}(\omega) \to 0$ as $\omega \to 0$ and

$$|U(j\omega)| \le l_{c1}(\omega) \qquad |\hat{P}^*(j\omega)^{-1}| \le l_{c2}(\omega) \tag{5.48}$$

for all $\omega \ge 0$.

(E3) The reference input $r(t)$ is stationary and $r(t) \in L_\infty$.

Remark: Assumption (E2) implies that

$$|\hat{P}(j\omega)^{-1}| \le \frac{l_{c2}(\omega)}{1-l_{c1}(\omega)} \tag{5.49}$$

for all $0 \le \omega \le \omega^0$ and some $\omega^0 > 0$.

Denote $H(\theta,s)$ to be the closed loop plant transfer function which is obtained by keeping the parameter $\theta(t)$ fixed at some θ. Similarly, let $n(\theta,s): R^{2n} \times C \to C^{2n}$ denote the column transfer function from the reference input $r(t)$ to the signal vector $w_\theta(t)$ defined by

$$w_\theta(t) = (r(t), v_\theta^{(1)}(t)^T, y_{p\theta}(t), v_\theta^{(2)}(t)^T) \tag{5.50}$$

where $v_\theta^{(1)}(t)$, $y_{p\theta}(t)$ and $v_\theta^{(2)}(t)$ are time functions obtained with $\theta(t)$ being fixed at θ. Then, referring to the structure of MRAC in Fig. 2.2, we can express $n(\theta,s)$ in terms of (Λ,b), $\hat{P}$ and $H(\theta,s)$ as:

$$n(\theta,s) = \begin{bmatrix} 1 \\ (sI-\Lambda)^{-1}b\hat{P}^{-1}(s)H(\theta,s) \\ H(\theta,s) \\ (sI-\Lambda)^{-1}bH(\theta,s) \end{bmatrix} = \begin{bmatrix} 1 \\ \overline{v}_{\hat{P}}(s)H(\theta,s) \end{bmatrix} \tag{5.51}$$

where $\overline{v}_{\hat{P}}(s):C\to C^{2n-1}$. Assuming that $H(\theta,s)$ is *stable*, we have

$$y_{p\theta} = \hat{P}(s)(\theta^T w_\theta) \tag{5.52}$$

and thus

$$\hat{P}^{-1}(s)H(\theta,s) = \theta^T n(\theta,s) = c_0+\theta_s^T\overline{v}_{\hat{P}}(s)H(\theta,s) \tag{5.53}$$

where

$$\theta_s^T = (c^T,d_0,d^T) \tag{5.54}$$

The closed loop plant transfer function $H(\theta,s)$ can thus be represented explicitly by θ as

$$H(\theta,s) = \frac{c_0}{\hat{P}^{-1}(s)-\theta_s^T\overline{v}_{\hat{P}}(s)} \tag{5.55}$$

Now suppose that the reference input $r(t)$ contains γ spectral lines $j\omega_1$, ..., $j\omega_\gamma$ such that the closed loop plant transfer function matches the model transfer function at these frequencies, i.e.

$$\hat{M}(j\omega_i) = H(\theta,j\omega_i) = \frac{c_0}{\hat{P}^{-1}(j\omega_i)-\theta_s^T\overline{v}_{\hat{P}}(j\omega_i)} \qquad i=1,\dots,\gamma \tag{5.56}$$

Then the parameter $\theta\in R^{2n}$ satisfies (5.54) if and only if it satisfies the following

$$\theta^T \begin{bmatrix} 1 \\ \overline{v}_{\hat{P}}(j\omega_i)\hat{M}(j\omega_i) \end{bmatrix} = \hat{P}^{-1}(j\omega_i)\hat{M}(j\omega_i) \qquad i=1,\dots,\gamma \tag{5.57}$$

or more concisely

$$\theta^T v_{\hat{P}}(j\omega_i) = \hat{P}^{-1}(j\omega_i)\hat{M}(j\omega_i) \qquad i=1,\dots,\gamma \tag{5.58}$$

(Notice that $\overline{n}(s) = v_{P^*}(s)$ where $\overline{n}(s)$ is defined in section 4) Grouping the set of equations in (5.58), we obtain a more compact form

$$\theta^T Y_\gamma = g_\gamma^* \tag{5.59}$$

where

$$Y_\gamma := \left[v_{\hat{P}}(j\omega_1), v_{\hat{P}}((j\omega_2),\cdots,v_{\hat{P}}(j\omega_{\gamma-1}),v_{\hat{P}}(j\omega_\gamma)\right] \tag{5.60}$$

and

$$g_\gamma^\bullet := \left[\hat{P}^{-1}(j\omega_1)\hat{M}(j\omega_1),\hat{P}^{-1}(j\omega_2)\hat{M}(j\omega_2),\cdots,\hat{P}^{-1}(j\omega_{\gamma-1})\hat{M}(j\omega_{\gamma-1}),\hat{P}^{-1}(j\omega_\gamma)\hat{M}(j\omega_\gamma)\right] \tag{5.61}$$

Note that

$$v_\beta^T(s) = \left[1,\frac{d_p(s)}{n_p(s)d_m(s)},\cdots,\frac{d_p(s)s^{n-2}}{n_p(s)d_m(s)},\frac{n_m(s)}{d_m(s)},\frac{1}{d_m(s)},\cdots,\frac{s^{n-2}}{d_m(s)}\right]$$

$$= \frac{1}{n_p(s)d_m(s)}\left[n_p(s)d_m(s),d_p(s),\cdots,d_p(s)s^{n-2},n_p(s)n_m(s),n_p(s),\cdots,n_p(s)s^{n-2}\right] \tag{5.62}$$

which has the similar structure to (5.7). Hence it may be verified ([12]) that Y_γ is of full row rank for almost all $\gamma\geq 2n$ spectral lines. In the following, we introduce a similar technical assumption to the one (D4) so that the case where a reference input consists of sinusoids can be dealt with.

Assumption:

(E4) There exist a set of frequencies $(\omega_1,\cdots,\omega_n,-\omega_1,\cdots,-\omega_n)$ such that the matrix Y_{2n} defined by

$$Y_{2n} := \left[v_\beta(j\omega_1),v_\beta((-j\omega_1),\cdots,v_\beta(j\omega_n),v_\beta(-j\omega_n)\right] \tag{5.63}$$

is nonsingular.

Based on this assumption, we give the following proposition.

Proposition 5.5:

Given the transfer function $v_\beta(s)$ as in (5.62) satisfying the assumption (E4). Define the matrix $Y_{2l}\in C^{2n\times 2l}$ as

$$Y_{2l} := \left[v_\beta(j\omega_1),v_\beta((-j\omega_1),\cdots,v_\beta(j\omega_l),v_\beta(-j\omega_l)\right] \tag{5.64}$$

Then for almost all $(\omega_1,\omega_2,\cdots,\omega_l)^T\in R^l$, $l\geq n$, the matrix Y_{2l} is of full row rank, i.e. $\rho(Y_{2l})=2n$. Equivalently, the set $\{\ (\omega_1,\omega_2,\cdots,\omega_l)^T\in R^l \mid \rho(Y_{2l})<2n\ \}$ is of zero measure.

Proof: cf. [12]

Corollary 5.6: (Almost Persistency of Excitation)

Consider the signal vector w_{θ_0} as defined in (5.50) with the assumption (E4) being satisfied. Then w_{θ_0} is PE for almost all reference inputs consisting of l sinusoids where $l\geq n$.

The proof can be easily obtained by using the transfer function $n(\theta_0,s)$ and proposition 5.5, and hence is omitted here.

Repeat the set of equations (5.59) here

$$\theta^T Y_{2l} = g_{2l}^* \tag{5.65}$$

It is obvious that, when $l \geq n$, there may not be a $\theta \in R^{2n}$ that solves (5.65). Instead, the minimum error (least square error) solutions, either unweighted or weighted , can be obtained and are given respectively by

$$\theta_0 = (Y_{2l}Y_{2l}^*)^{-1}Y_{2l}g_{2l} \tag{5.66}$$

and

$$\theta_0 = (Y_{2l}WY_{2l}^*)^{-1}Y_{2l}Wg_{2l} \tag{5.67}$$

where W is a weighting matrix similarly defined as in (5.29) and (5.25). This parameter θ_0 will be called *tuned parameter* if the resulting transfer function $H(\theta_0,s)$ is a stable transfer function and $H(\theta_0,s)$ will then be called *tuned plant* transfer function.

With the tuned plant notion, it is shown in [10] that the MRAC system can be transformed into the following

$$\dot{e} = (A_{\theta_0}+b\phi^T Q)e+bw_{\theta_0}^T\phi \tag{5.68}$$

$$\dot{\phi} = -\varepsilon(w_{\theta_0}h^T+e_{\theta_0}Q)e-\varepsilon Qeh^Te-\varepsilon e_{\theta_0}w_{\theta_0} \tag{5.69}$$

with e being defined as the state error between the states in the closed loop plant and the tuned plant and ϕ as $\theta-\theta_0$, where

$$e_{\theta_0} := y_{p\theta_0} - y_m = H(\theta_0,s)(r) - \hat{M}(s)(r) \tag{5.70}$$

is called the *tuned error*, A_{θ_0} is a Hurwitz matrix,

$$c_0^o h^T(sI-A_{\theta_0})^{-1}b = H(\theta_0,s)$$

with $\theta_0^T = (c_0^o, c^{oT}, d_0^o, d^{oT})$, and Q is the matrix defined in section 2. In the following, we will discuss two cases where the reference input consists of n sinusoids and more than n sinusoids. We will use the corollary 5.6 to assume that w_{θ_0} is PE and leave out the "almost-allness" temporarily.

Case I: $l=n$

By definition, $H(\theta_0,s)$ is a stable transfer function and the frequency response of the tuned error $\hat{e}_{\theta_0}(j\omega_i)=0$ for i = -n,..,-1,1,..,2 where $j\omega_i$ is the the frequency of the ith sinusoid in the reference input. Hence, the tuned error $e_{\theta_0}(t)$ is an exponentially decaying function, i.e. $e_{\theta_0}(t)\to 0$ exponentially as $t\to\infty$. Moreover, when the averaging assumptions are met, the asymptotic behavior of the system (5.68) (5.69) is equivalent to the following system

$$\dot{e} = (A_{\theta_0}+b\phi^T Q)e+bw_{\theta_0}^T\phi \tag{5.71}$$

$$\dot{\phi} = -\,\varepsilon(w_{\theta_0}h^T+e_{\theta_0}Q)e-\,\varepsilon Qeh^Te \tag{5.72}$$

which is shown ([10],[4]) to be exponentially stable. Consequently, the parameter error of the original system (5.68) (5.69), $\phi=\theta-\theta_0$, satisfies $\phi\to 0$ as $t\to\infty$.

CASE II: $l>n$

The tuned error $e_{\theta_0}(t)$ in this case is in general a nonzero bounded signal. To prevent from the *slow-drift instability* in the presence of unmodelled dynamics, we further require that the reference input $r(t)$ be a *dominantly rich good signal* as defined in the work [10]. Now if $|\hat{e}_{\theta_0}(j\omega_i)|\le g_b'$ for i = $-l$,..,-1,1,...,l, and hence $|e_{\theta_0}(t)| \le g_b$, it is shown ([13]) that there exist some class k function of ε, $\overline{\psi}(\varepsilon)$, and a nondecreasing function $\delta(a){:}R_+\to R_+$ and $\delta(a)\to 0$ as $a\to 0$ such that $\phi(t)$ is a bounded function of $t>0$ and

$$\lim_{t\to\infty} \|\phi(t)\| \le \delta(g_b)+\overline{\psi}(\varepsilon) \tag{5.73}$$

provided $\|e(0)\|$, $\|\phi(0)\|$, g_b, and ε are small enough.

Remark: In general, the way to choose a stationary input $r(t)$ with finite number of frequencies such that $H(\theta_0,s)$ (θ_0 as defined in (5.67)) is a tuned plant transfer function and, incidentally, that $|\hat{e}_{\theta_0}(s)|$ is small enough, is to select frequencies in the midband region of the model transfer function. Also, it is more possible to have so obtained input be a dominantly rich good signal as required to prevent from slow-drift instability.

Next, the bound on the difference between the weighted tuned parameter θ_0 defined in (5.67) and the nominal parameter θ^* corresponding to the nominal plant, namely, $\|\theta_0-\theta^*\|$, can be evaluated in terms of the frequency content of the input and the plant uncertainty. This result is then used to derive a bound on the frequency search range Ω required in the numerical design algorithm.

Proposition 5.7: (Bound on Ω in MRAC)

Consider a Model Reference Adaptive Controller satisfying assumptions (E1)-(E4). Suppose now the reference input $r(t)$ is power constrained and contains finite number, say $s \geq n$, of sinusoids with frequencies lying in the compact set $\Omega = [-\overline{\omega}, \overline{\omega}]$ in the frequency spectrum, i.e. we have

$$\frac{1}{2\pi} \int_{\Omega} dF_r(\omega) = \sum_{i=-s}^{s} \frac{a_i^2}{t_i} = F_\infty \qquad a_0 = 0 \tag{5.74}$$

where a_i is similarly defined as in (5.21). If $\overline{\omega}$ satisfies

$$\overline{\omega} \leq \omega^0 \quad \text{and} \quad 0 < \frac{\alpha_4(\overline{\omega})}{(\, \lambda[R_{w_m}(0)] - \alpha_5(\overline{\omega}) \,)} \leq \eta \tag{5.75}$$

for some small $\eta > 0$ where $\alpha_4(\cdot)$, $\alpha_5(\cdot)$ are some positive nondecreasing functions and $\alpha_4(a)$, $\alpha_5(a) \to 0$ as $a \to 0$, then

$$\|\theta_0 - \theta^*\| \leq \eta \|\theta^*\|$$

Proof: See Appendix.

Remark: Similar to the case of adaptive identification, $\overline{\omega}$ can be derived numerically from (5.75) and thus determines the frequency search range Ω which is used in the numerical input design algorithm.

Example 2: Consider the second example in section 4 with the plant contaminated by a high frequency unmodelled pole $s=-20$, i.e.

$$\hat{P}(s) = \frac{2}{(s+1)} \frac{20}{(s+20)} \tag{5.76}$$

Recall that the nominal parameter θ^* is: $k_r^* = 1.5$ and $k_y^* = -1$, and the optimal input design obtained in the second example in section 4 contains only single frequency. From previous discussions, the adjustable parameter θ then converges to the tuned parameter θ_0, where $k_r = 1.575$ and $k_y = -1.226$, which gives the tuned plant transfer function

$$\hat{P}_{\theta_0}(s) = \frac{63}{(s^2 + 21s + 69.05)} \tag{5.77}$$

Fig. 5.4 (a) (b) (c) demonstrate the time trajectories of the parameter error ϕ_1, ϕ_2 and the output error $e = y_p - y_m$ using the initial design F^0 as in (4.49); whereas Fig. 5.5 (a) (b) (c) show the time trajectories for the same parameter error and the output error using the optimal input design. The rate of convergence of the parameter in the case where the optimal input is applied is almost twice faster than that in the case where the non-optimal input is used. In Fig. 5.6, we draw the Nyquist plots for the transfer functions of the model and the tuned plant to show the closeness between them.

Conclusion:

In this paper, we propose a frequency-domain input design algorithm and apply it to both Adaptive Identification and Model Reference Adaptive Control schemes for continuous time case. It is not hard to show that, by using the formulation of the discrete time case in [11] [15], the same design algorithm can be applied to the discrete time adaptive system as well. We feel that the algorithm presented here will become a very useful design tool if a good estimate of the unknown plant can be obtained beforehand.

Acknowledgements

We would like to thank Erwei Bai for several very useful discussions. Earley comments from Marc Bodson and Jeff Mason are acknowledged. We also like to thank Andy Packard and Jeff Mason for their helpful proofreading of this paper.

This research has been supported by NASA under grant NAG 2-243 and Army Research Office under grant DAAG-29-85-k-0072

References:

[1] K. J. Astrom, "Interactions Between Excitation and Unmodelled Dynamics," IEEE Trans. on AC, Vol. AC-30, No. 9, Sep. 1985, pp 889-891

[2] E. W. Bai and S. S. Sastry, "Persistency of Excitation, Sufficient Richness and Parameter Convergence in Discrete Time Adaptive Control," System and Control Letters, Vol. 6, (1985), pp 153-163

[3] Marc Bodson and S. S. Sastry, "Small Signal I/O Stability of Nonlinear Control Systems: Application to the Robustness of a MRAC Scheme," Memorandum No. UCB/ERL M84/70, Electronics Research Laboratory, University of California, Berkeley, 1984

[4] Marc Bodson, S. S. Sastry, B.D.O. Anderson, I. Mareels and R. Bitmead, "Nonlinear Averaging Theorems, and the Determination of Parameter Convergence Rates in Adaptive Control," System and Control Letters 7, 1986, pp. 145-157

[5] Marc Bodson and S. S. Sastry, "Stability, Convergence, and Robustness of Adaptive Systems," Memorandum No. UCB/ERL M86/66, 25 August, 1986, Electronics Research Laboratory, University of California, Berkeley

[6] S. Boyd and S. S. Sastry, "Necessary and Sufficient Conidtions for Parameter Convergence in Adaptive Control," Proc. of the Ames-Berkeley Conf. on Non-Linear Dynamics and Control, Math-Sci Press, Brooklin, Massachusetts, 1984, pp 81-101

[7] S. Boyd and S. S. Sastry, "On Parameter Convergence in Adaptive Control," System and Control Letters, Vol. 3, 1983, pp 311-319

[8] A. S. Deif, "Advanced Matrix Theory for Scientists and Engineers," Abacus Press, Tunbridge Wells & London, Halsted Press Division John Wiley & Sons, New York & Toronto

[9] Li-Chen Fu, M. Bodson and S. S. Sastry, "New Stability Theorems for Averaging and Their Application to Convergence Analysis of Adaptive Identification and Control Schemes," Proc. of 24th IEEE Conf. on Decision and Control, Ft. Lauderdale, FL, 1985, pp 473-477

[10] Li-Chen Fu and S. S. Sastry, "Slow Drift Instability in Model Reference Adaptive System -- An Averaging Analysis," to appear in Intl. J. of Control, Jan. 1987

[11] Li-Chen Fu, E. W. Bai and S. S. Sastry, "Averaging Analysis for Discrete Time and Sampled Data Adaptive Systems," Proc. of 25th IEEE Conf. on Decision and Control, Athens, Greece, 1986

[12] E. W. Bai, Marc Bodson, L. C. Fu, J. Mason, and S. S. Sastry, "Adaptive Identification in the Presence of Unmodelled Dynamics," in preparation, preprint will be ready in Feb.

1987

[13] Li-Chen Fu and S. S. Sastry, "Averaging Analysis of Model Reference Adaptive Control with Unmodelled Dynamics," in preparation, preprint will be ready in April 1987

[14] V. V. Fedorov, "Theory of Optimal Experiments," Academic Press, New York, 1972

[15] G. C. Goodwin and K. S. Sin, "Adaptive Filtering Prediction and Control," Prentice-Hall, Inc., Englewood Cliffs, New Jersey 07632, 1984

[16] G. C. Goodwin and R. L. Payne, "Dynamic System Identification," Academic Press, New York, 1977

[17] P.A. Ioannou and P.V. Kokotovic, "Robust Redesign of Adaptive Control," IEEE Trans. on AC, Vol. AC-29, No. 3, pp 202-211, 1984

[18] P.A. Ioannou and K.S. Tsakalis, "A Robust Direct Adaptive Controller," IEEE Trans. on AC, Vol. AC-31, No. 11, pp 1033-1043, 1986

[19] G. Kreisselmeier, "Adaptive Observers with Exponential Rate of Convergence," IEEE Trans. on AC, Vol. AC-22, pp 2-8, 1977

[20] G. Kreisselmeier and K.S. Narendra, "Stable Model Reference Adaptive Control, "IEEE Trans. on AC, Vol. AC-27, No. 6, pp 1169-1175, 1982

[21] R. Kosut, B.D.O. Anderson and I. Mareels, "Stability Theory for Adaptive Systems: Method of Averaging and Persistency of Excitation," IEEE Trans. on AC, Vol. AC-32, pp 26-34, 1986

[22] T. Kato, "Perturbation Theory for Linear Operator," Springer-Verlag, New York, 1980

[23] G. Lunders and K. S. Narendra, "An Adaptive Observer and Identifier for a Linear System," IEEE Trans. on AC, Vol. AC-18, pp 496-499, 1973

[24] I.M.Y. Mareels, R. Bitmead, M. Gevers, C.R. Johnson, R.L. Kosut and M.A. Poubelle, "How Exciting Can a Signal Really Be?," to appear in System and Control Letters

[25] R. K. Mehra, "Frequency-Domain Synthesis of Optimal Inputs for Linear System Parameter Estimation," Trans. of ASME, June 1976, pp 130-138

[26] R. K. Mehra, "Optimal Input Signals for Parameter Estimation in Dynamic Systems--Survey and New Results, IEEE Trans. on AC, Vol. AC-19, pp 753-763, 1974

[27] K. Narendra and L. Valavani, "Stable Adaptive Controller Design- Direct Controller," IEEE Trans. on AC, Vol. AC-23, pp 570-583, 1978

[28] K. Narendra, Y.-H. Lin and L. Valavani, "Stable Adaptive Controller Design, Part II: Proof of Stability," IEEE Trans. on AC, Vol. AC-25, No. 3, pp 440-448, 1980

[29] B.B. Peterson and K.S. Narendra, "Bounded Error Adaptive Control," IEEE Trans. on AC, Vol. AC-27, No. 6, pp 1161-1168, 1982

[30] C. Rohrs, L. Valavani, M. Athans and G. Stein, " Robustness of Continuous-Time Adaptive Control Algorithms in the Presence of Unmodelled Dynamics," IEEE Trans. on AC, Vol. AC-30, No. 9, pp 881-889, 1985

[31] M. Sondhi and D. Metra, "New Results on the Performance of a Well-Known Class of Adaptive Filters," Proc. of the IEEE, Vol. 64, No. 11, Nov. 1976

[32] C.K. Sanathanan and J. Koerner, "Transfer Function Synthesis as a Ratio of Two Complex Polynomials," IEEE Trans. on AC, Vol. AC-8, No. 1, pp 56-58, 1963

[33] S. S. Sastry, "Model-Reference Adaptive Control-Stability, Parameter Convergence, and Robustness, " IMA Journal of Mathematical Control & Information, Vol. 1, pp 27-66, 1984

[34] H. Stall, "Transfer Function Synthesis Using Frequency Response Data," Int. Journal of Control, Vol. 39, No. 3, pp 541-550, 1984

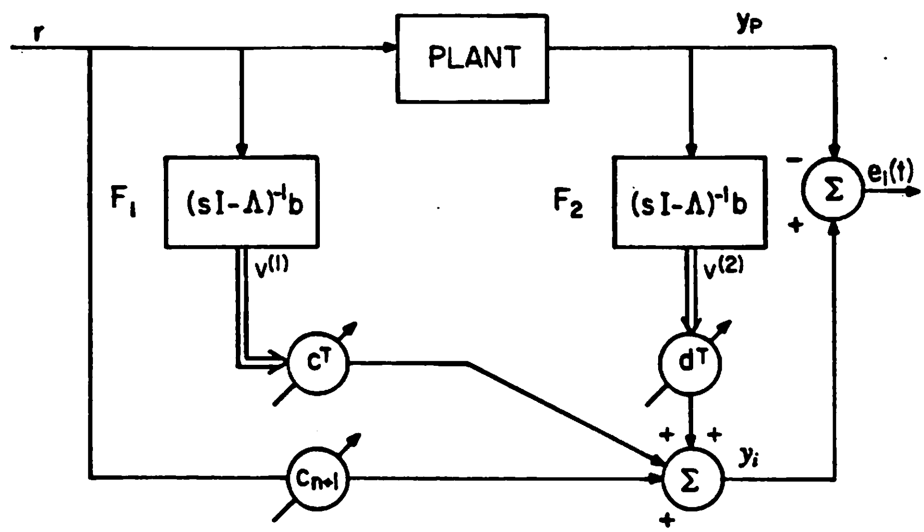

r
PLANT
y_P
F_1
$(sI-\Lambda)^{-1}b$
F_2
$(sI-\Lambda)^{-1}b$
$e_1(t)$
$v^{(1)}$
$v^{(2)}$
c^T
d^T
c_{n+1}
y_i

Fig. 2.1

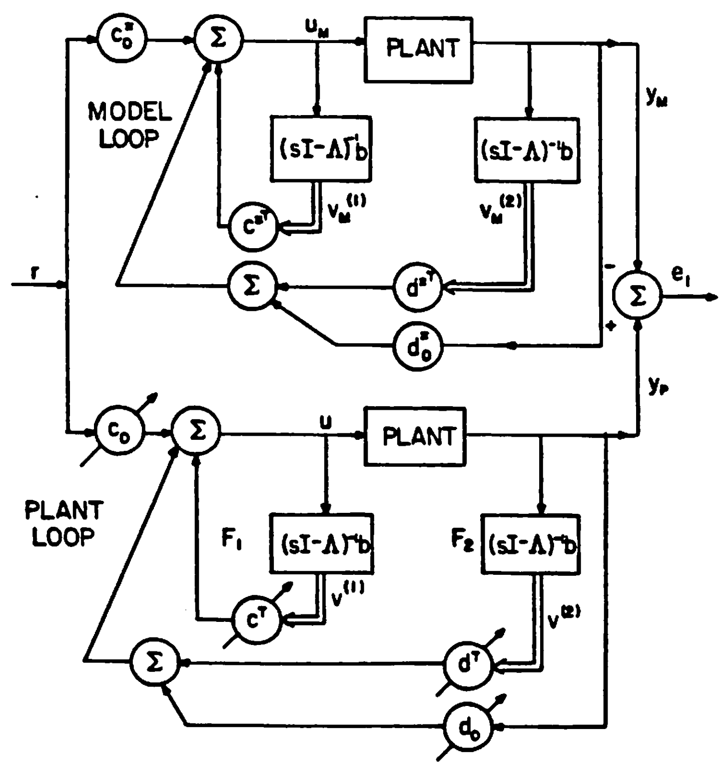

c_0^*
u_M
PLANT
MODEL LOOP
$(sI-\Lambda)^{-1}b$
$(sI-\Lambda)^{-1}b$
y_M
$v_M^{(1)}$
$v_M^{(2)}$
c^{*T}
r
d^{*T}
e_1
d_0^*
y_P
c_0
u
PLANT
PLANT LOOP
F_1
$(sI-\Lambda)^{-1}b$
F_2
$(sI-\Lambda)^{-1}b$
$v^{(1)}$
c^T
$v^{(2)}$
d^T
d_0

Fig. 2.2

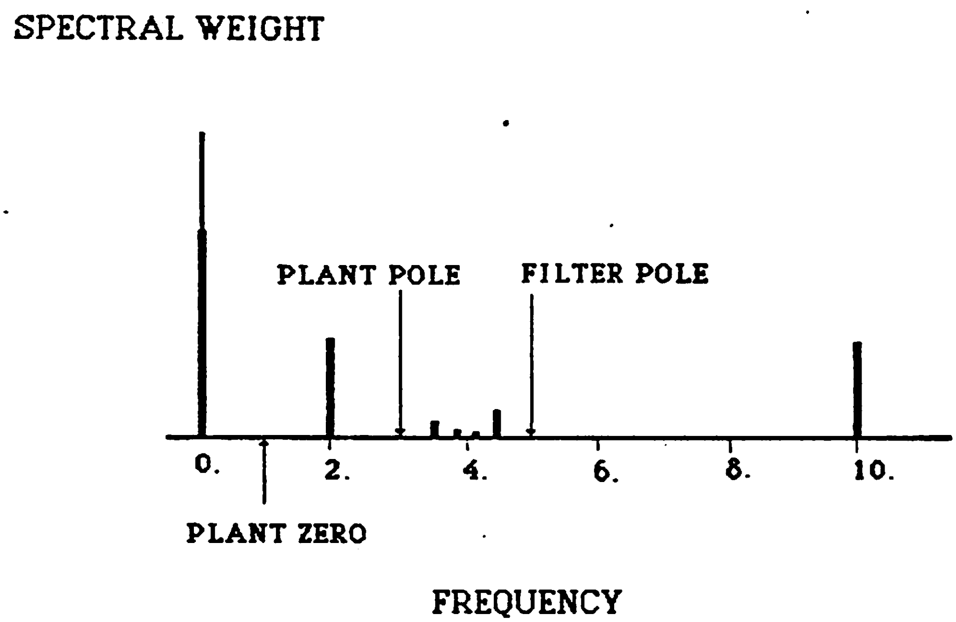

SPECTRAL WEIGHT
PLANT POLE
FILTER POLE
0.
2.
4.
6.
8.
10.
PLANT ZERO
FREQUENCY

Fig. 4.1

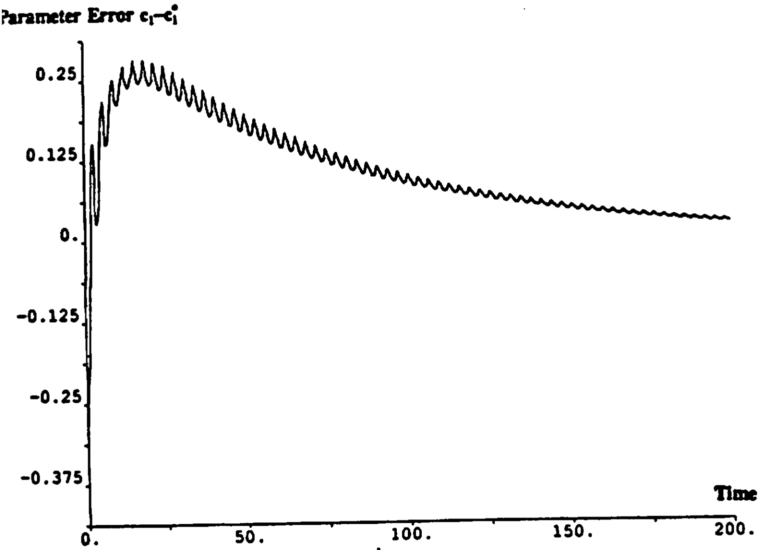


Fig. 4.2 (a)

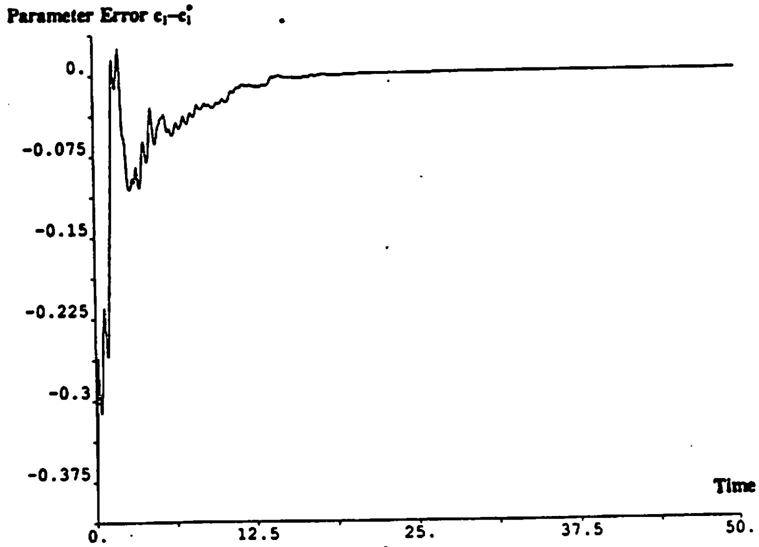


Fig. 4.3 (a)

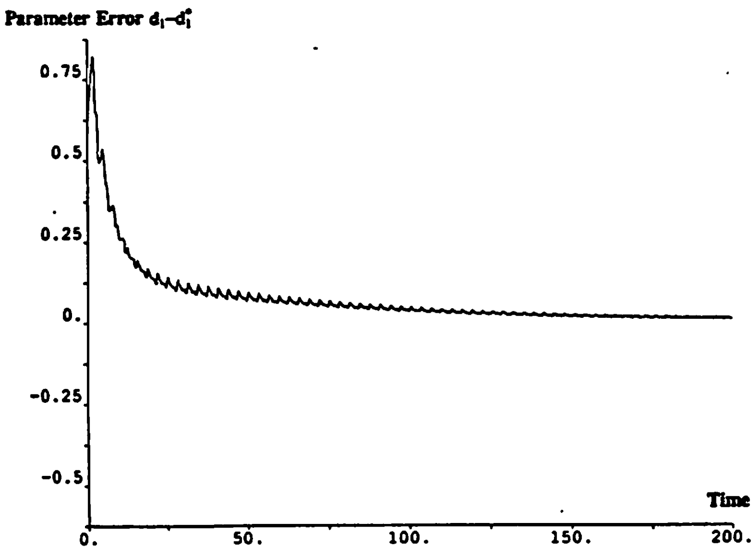


Fig. 4.2 (b)

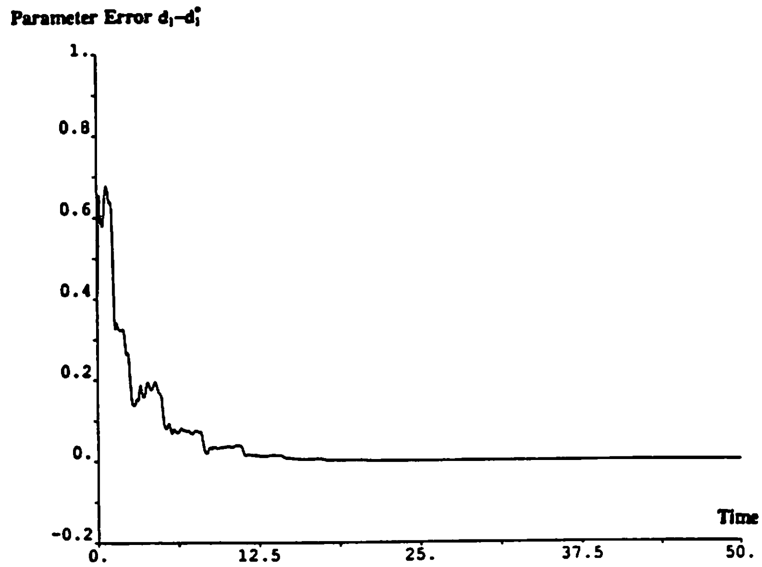


Fig. 4.3 (b)

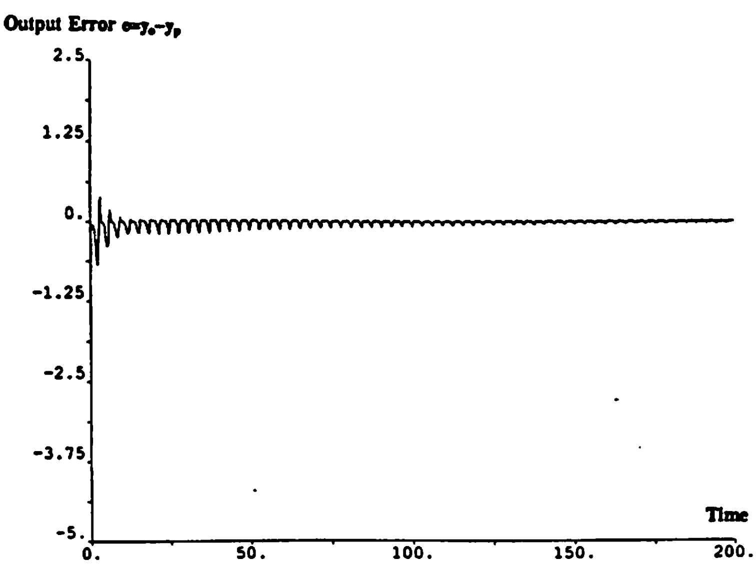


Fig. 4.2 (c)

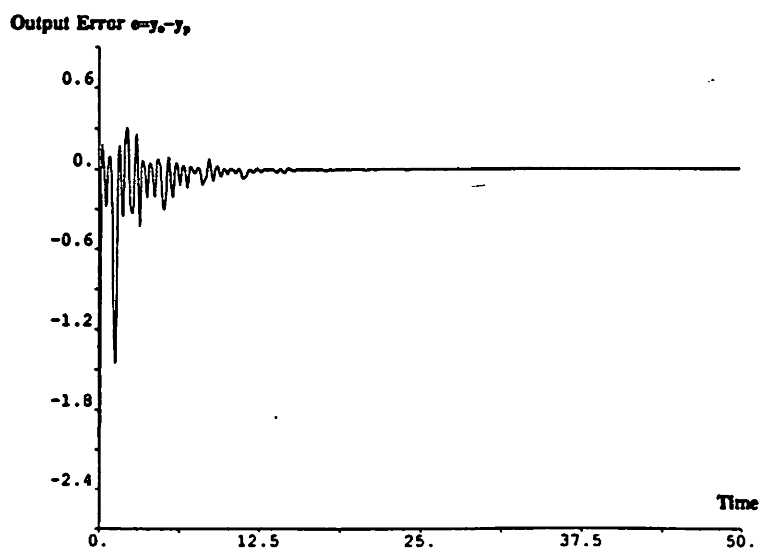


Fig. 4.3 (c)

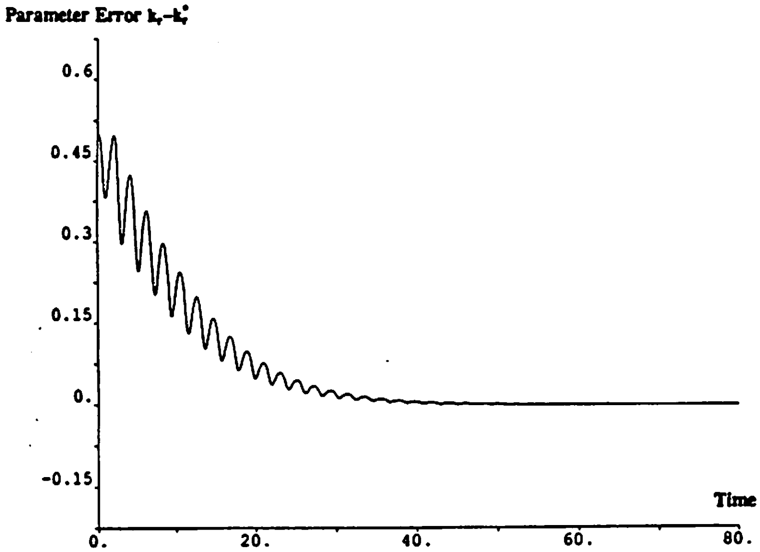


Fig. 4.4 (a)

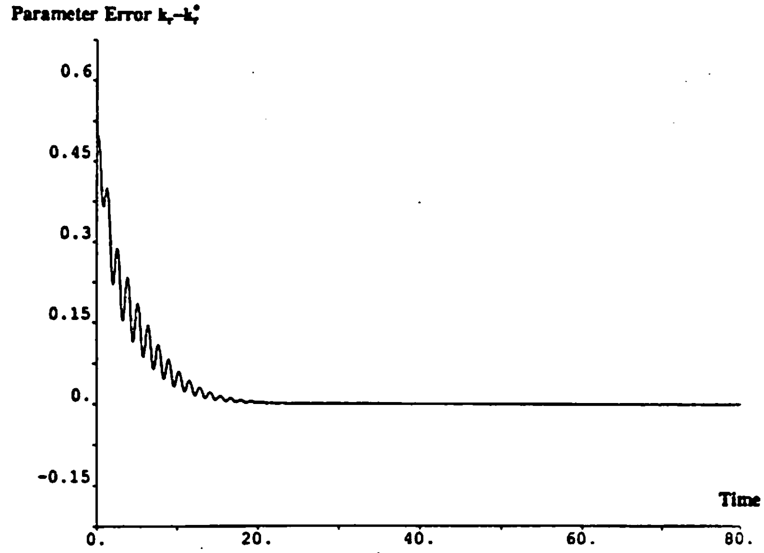


Fig. 4.5 (a)

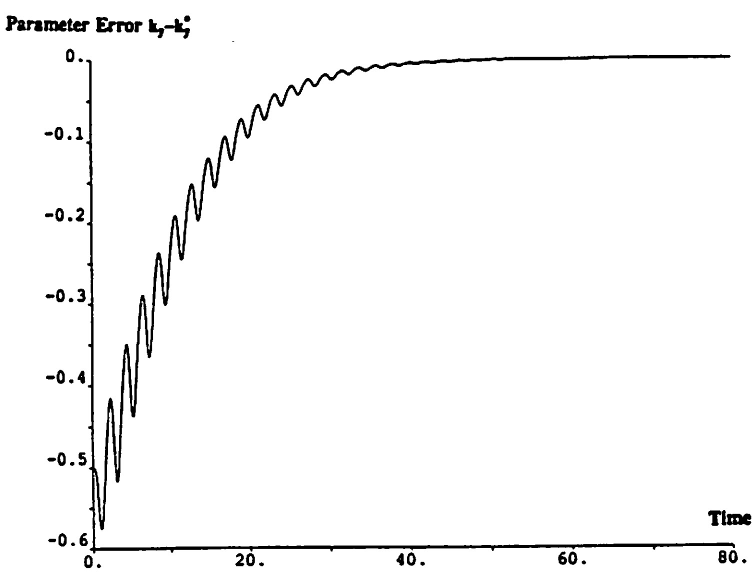


Fig. 4.4 (b)

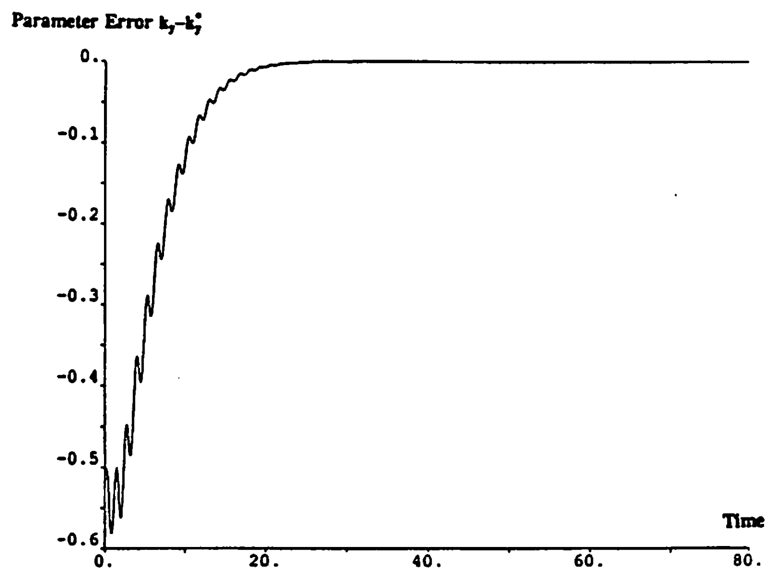


Fig. 4.5 (b)

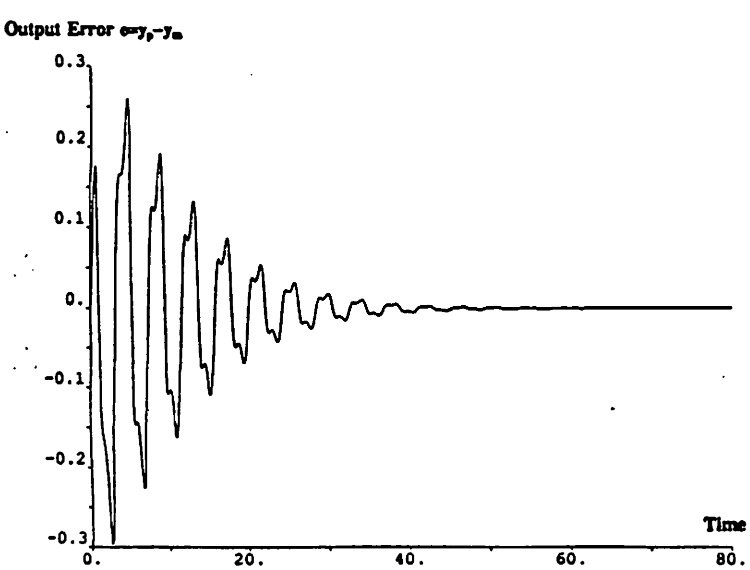


Fig. 4.4 (c)

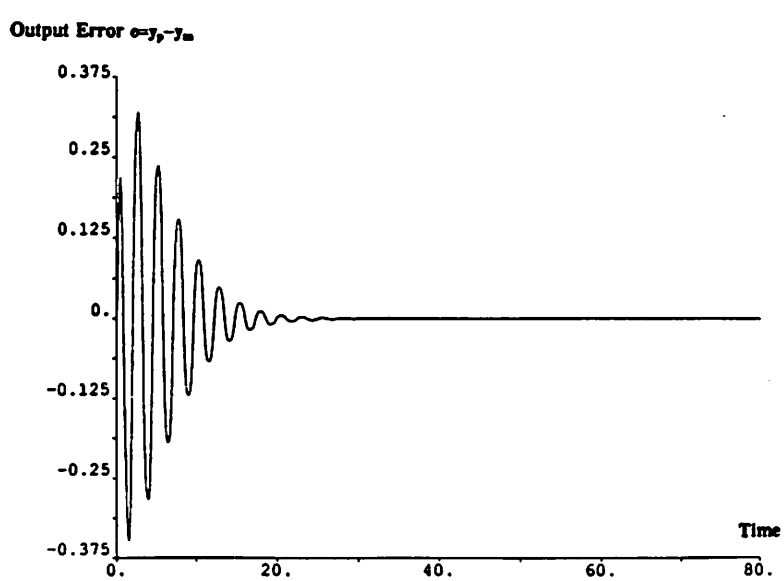


Fig. 4.5 (c)

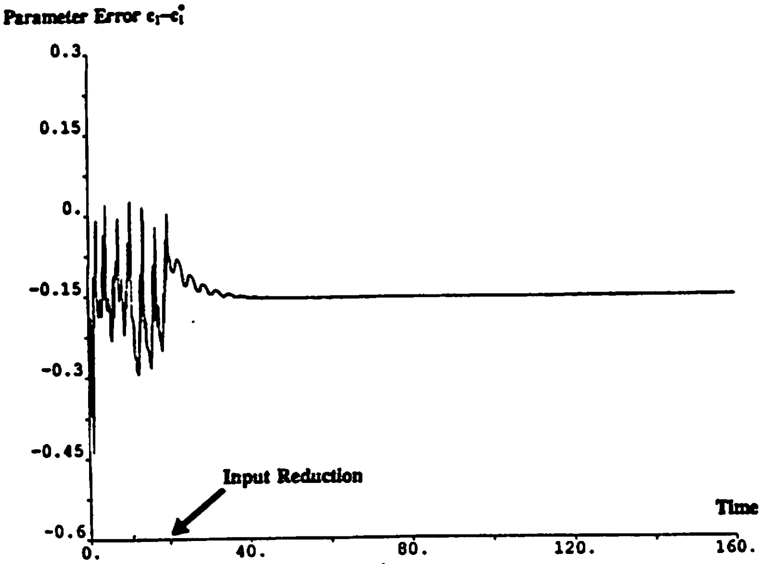


Fig. 5.1 (a)

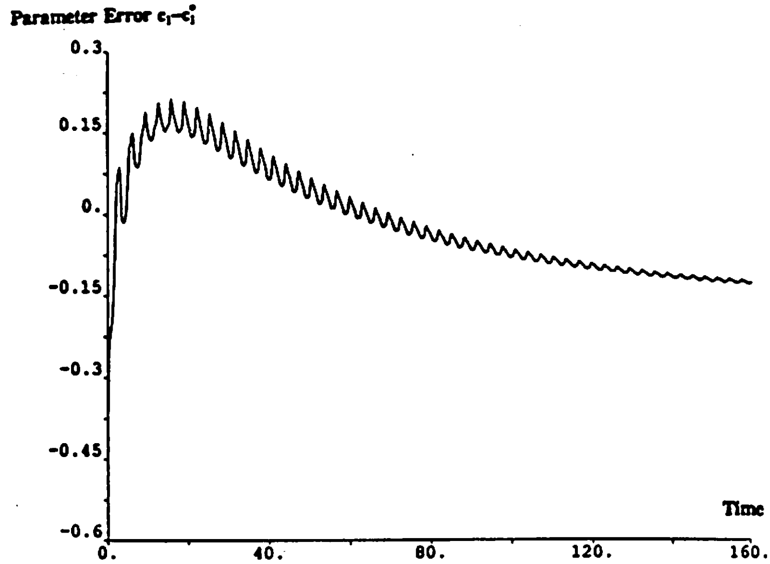


Fig. 5.2 (a)

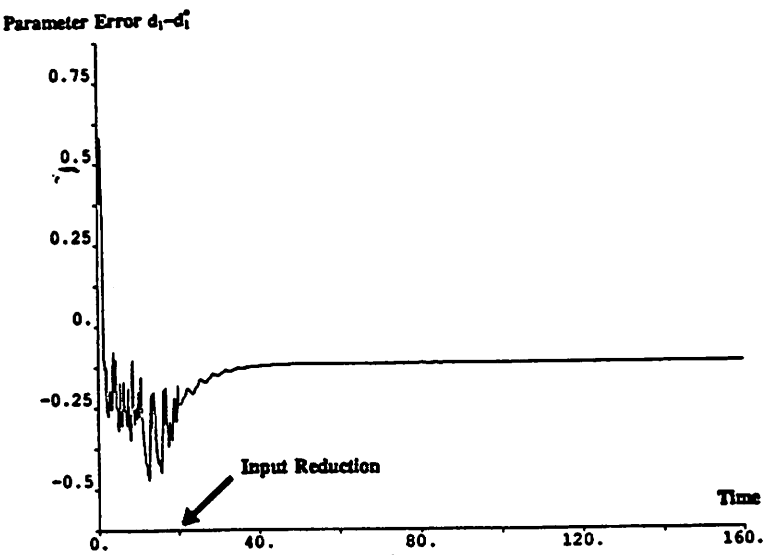


Fig. 5.1 (b)

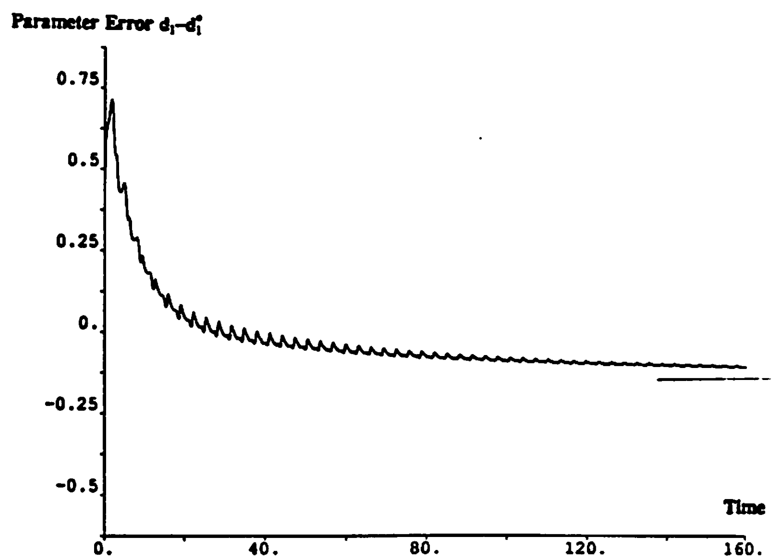


Fig. 5.2 (b)

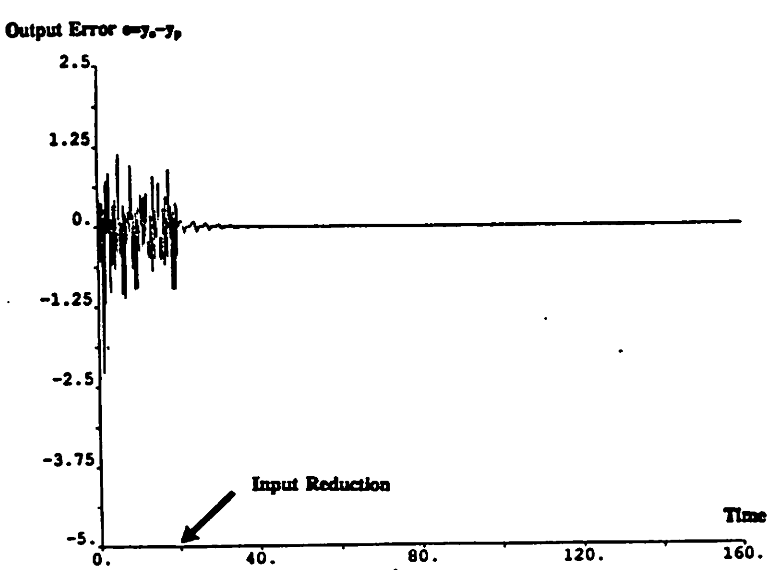


Fig. 5.1 (c)

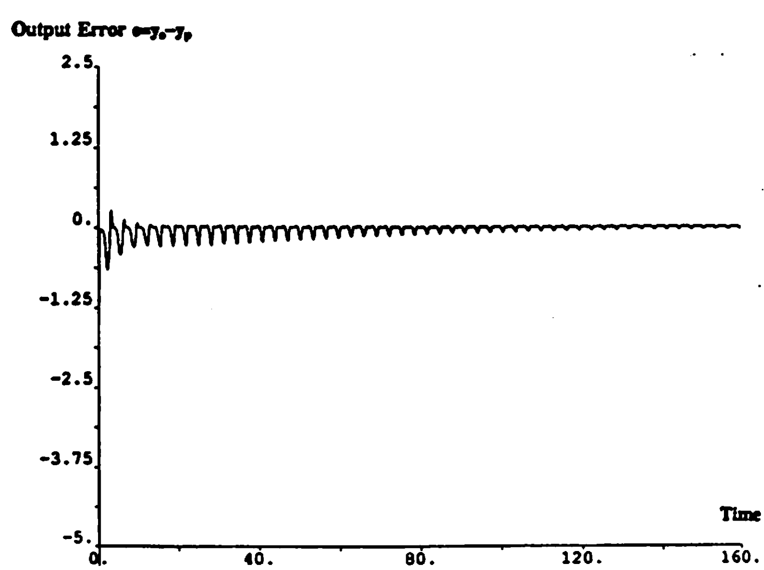


Fig. 5.2 (c)

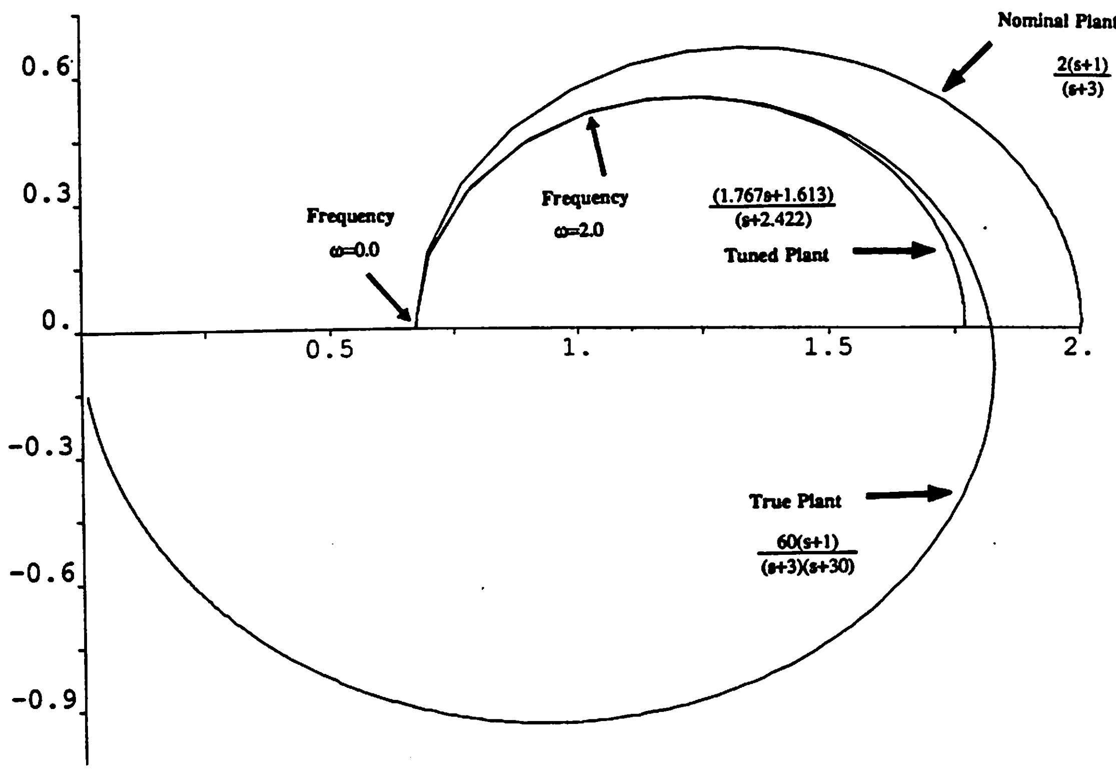

Nominal Plant
2(s+1) / (s+3)
0.6
0.3
0.
-0.3
-0.6
-0.9
0.5
1.
1.5
2.
Frequency
ω=0.0
Frequency
ω=2.0
(1.767s+1.613) / (s+2.422)
Tuned Plant
True Plant
60(s+1) / (s+3)(s+30)

Fig. 5.3

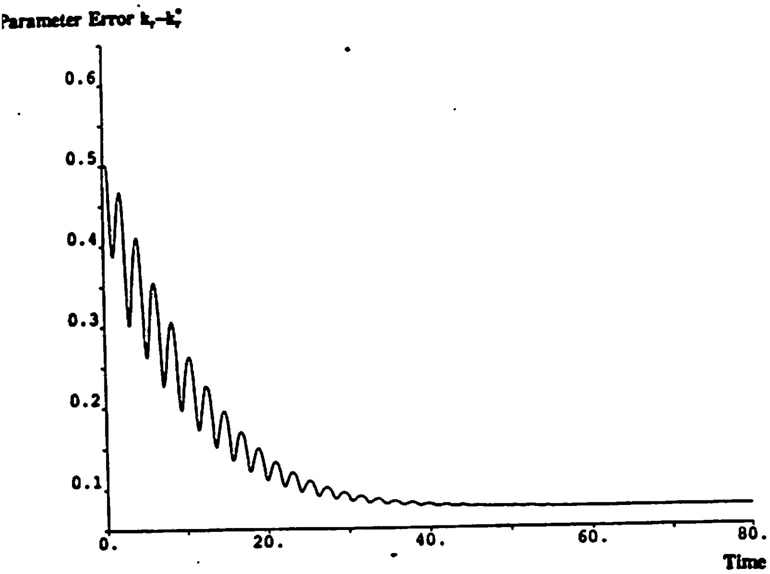


Fig. 5.4 (a)

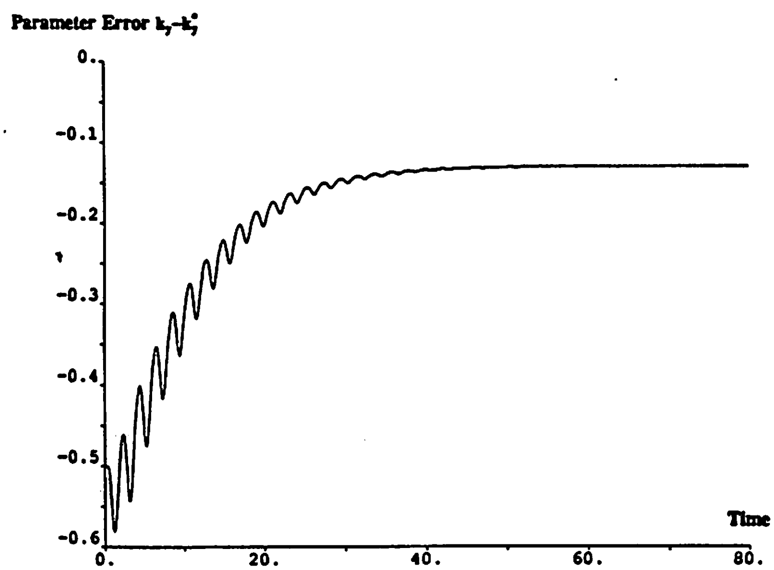


Fig. 5.4 (b)

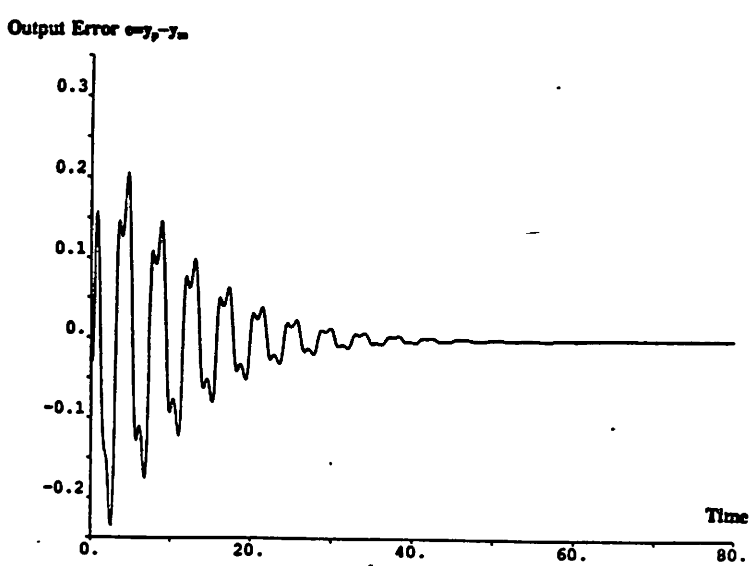


Fig. 5.4 (c)

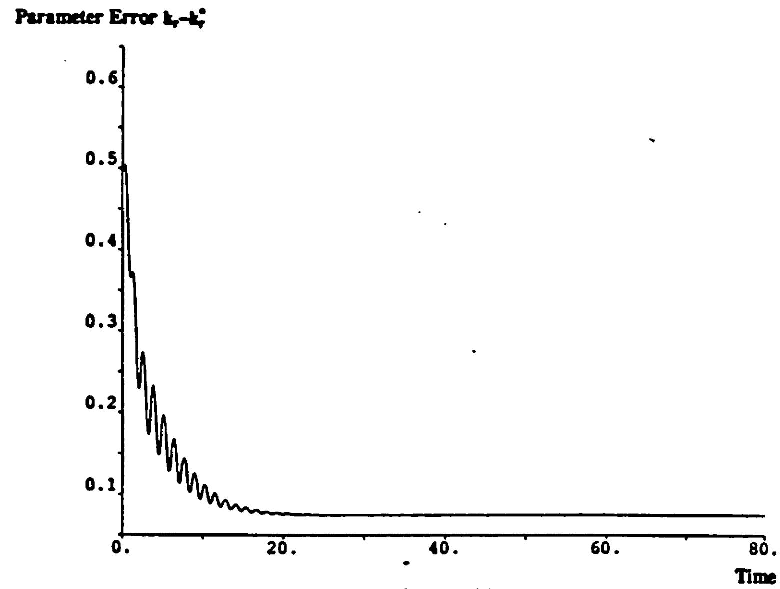


Fig. 5.5 (a)

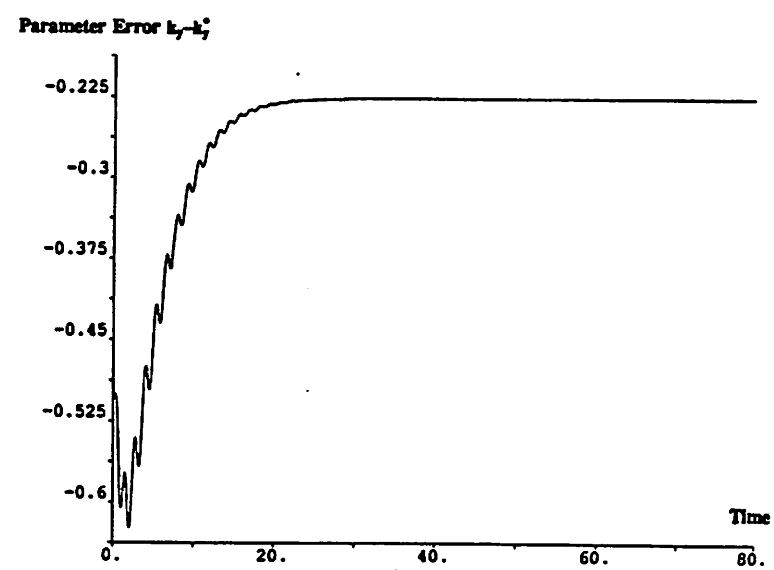


Fig. 5.5 (b)

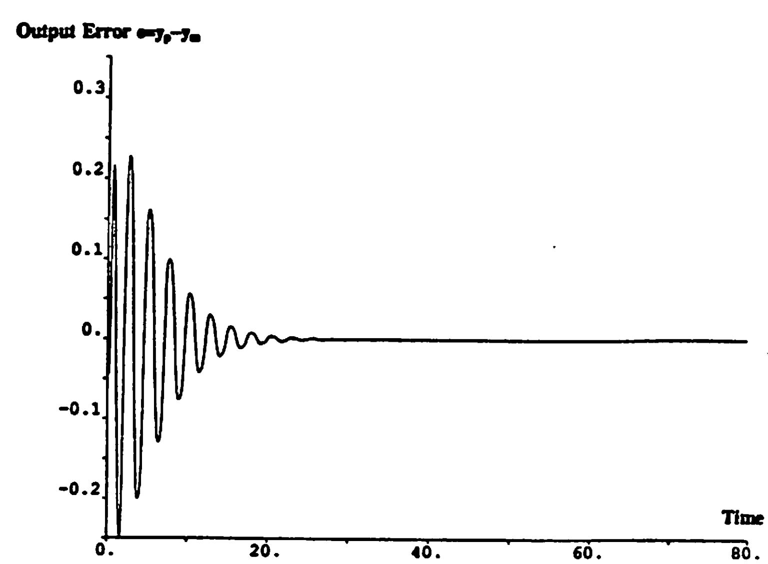
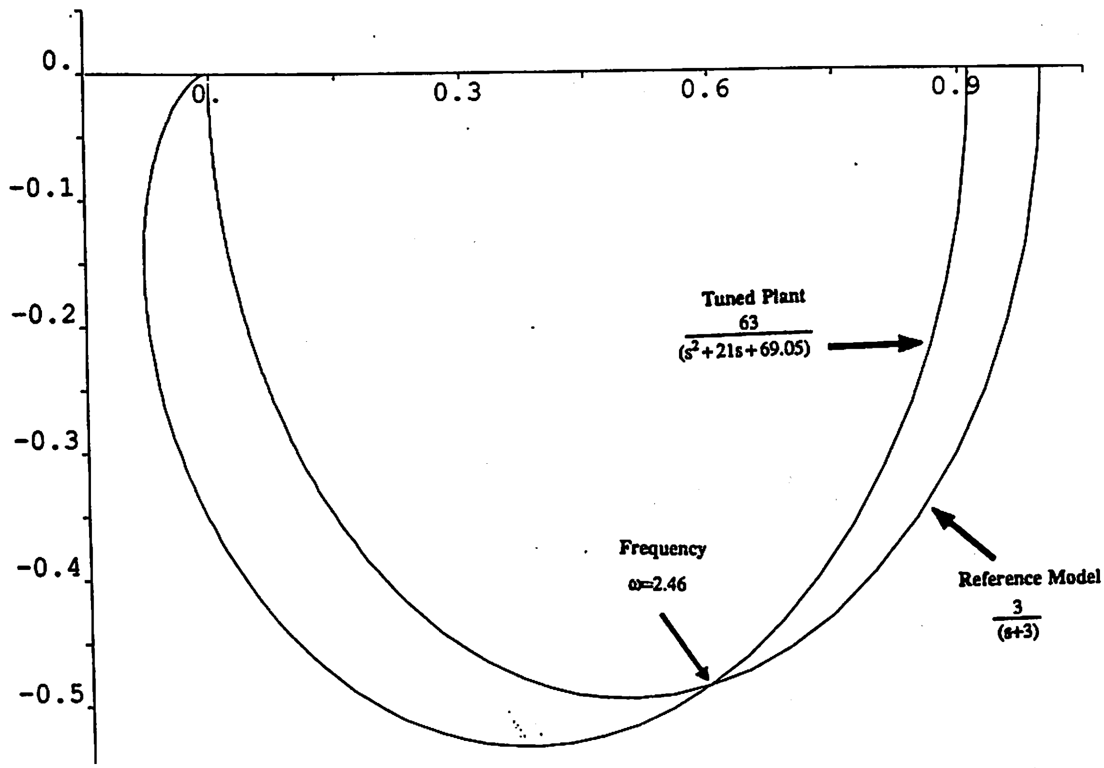


Fig. 5.5 (c)



0.
-0.1
-0.2
-0.3
-0.4
-0.5
0.
0.3
0.6
0.9
Tuned Plant
$\frac{63}{(s^2+21s+69.05)}$
Frequency
ω=2.46
Reference Model
$\frac{3}{(s+3)}$

Fig. 5.6

Appendix:

Proof of Proposition 5.1:

It can be seen from (5.7) that there exists a nonsingular matrix $T \in R^{(2n+1)\times(2n+1)}$ and a transfer function $n'(s)$ such that

$$n'(s) := Tn(s) = \frac{1}{\hat{\Lambda}(s)d_p(s)}\left[d_p(s)\,,\,sd_p(s)\,,\,\cdots\,,\,s^n d_p(s)\,,\,n_p(s)\,,\cdots,\,s^{n-1}n_p(s)\right]^T \tag{a.1}$$

and a matrix Z_n' defined by

$$Z_{2n+1}' := \left[n'(j\omega_1)\,,\,n'(j\omega_2)\,,\,\cdots\,,\,n'(j\omega_{2n})\,,\,n'(j\omega_{2n+1})\right] \tag{a.2}$$

is nonsingular if and only if Z_n is nonsingular since $Z_n'=TZ_n$. Similarly, we define $n_\gamma'(s)$ as

$$n_\gamma'(s) = \frac{1}{\hat{\Lambda}(s)d_p(s)}\left[d_p(s)\,,\,\cdots\,,\,s^{k_1}d_p(s)\,,\,n_p(s)\,,\,\cdots\,,\,s^{k_2}n_p(s)\right]^T \tag{a.3}$$

where $\gamma \le 2n+1$ and

$$\begin{cases} k_1=\dfrac{\gamma-2}{2} \quad k_2=\dfrac{\gamma-2}{2} & \gamma \text{ is even} \\ k_1=\dfrac{\gamma-1}{2} \quad k_2=\dfrac{\gamma-3}{2} & \gamma \text{ is odd} \end{cases}$$

Then it follows from the result of [12] that the matrix $\bar{Z}_\gamma \in C^{\gamma\times\gamma}$ defined as

$$\bar{Z}_\gamma := \left[n_\gamma'(j\omega_1)\,,\,n_\gamma'(j\omega_2)\,,\,\cdots\,,\,n_\gamma'(j\omega_{\gamma-1})\,,\,n_\gamma'(j\omega_\gamma)\right] \tag{a.4}$$

is nonsingular for almost all $(\omega_1,\cdots,\omega_\gamma)^T \in R^\gamma$. Furthermore, Z_γ can be related to $\bar{Z}_\gamma$ by

$$Z_\gamma = T^{-1}\,Z'_\gamma$$

$$= T^{-1}\begin{bmatrix} \dfrac{(j\omega_1)^n}{\hat{\Lambda}(j\omega_1)} & & \dfrac{(j\omega_\gamma)^n}{\hat{\Lambda}(j\omega\gamma)} \\ \vdots & & \vdots \\ \dfrac{(j\omega_1)^{k_1+1}}{\hat{\Lambda}(j\omega_1)} & \cdots & \dfrac{(j\omega_\gamma)^{k_1+1}}{\hat{\Lambda}(j\omega_\gamma)} \\ & \bar{Z}_\gamma & \\ \dfrac{n_p(j\omega_1)(j\omega_1)^{n-1}}{\hat{\Lambda}(j\omega_1)d_p(j\omega_1)} & \cdots & \dfrac{n_p(j\omega_\gamma)(j\omega_\gamma)^{n-1}}{\hat{\Lambda}(j\omega_\gamma)d_p(j\omega_\gamma)} \\ \vdots & & \vdots \\ \dfrac{n_p(j\omega_1)(j\omega_1)^{k_2+1}}{\hat{\Lambda}(j\omega_1)d_p(j\omega_1)} & & \dfrac{n_p(j\omega_\gamma)(j\omega_\gamma)^{k_2+1}}{\hat{\Lambda}(j\omega_\gamma)d_p(j\omega_\gamma)}. \end{bmatrix} \tag{a.5}$$

It then follows that Z_γ is of full column rank or $\rho(Z_\gamma)=\gamma$ for almost all $(\omega_1,\cdots,\omega_\gamma)^T \in R^\gamma$. In other words, the set $\{ (\omega_1,\cdots,\omega_\gamma)^T \in R^\gamma \mid \rho(Z_\gamma)<\gamma \}$ is of zero measure.

.Q.E.D.

Proof of Proposition 5.2:

(i) The proof can be approached by using the similar technique in [12].

(ii) Referring to (a.3), we have $n_{2l+1}'(s)$ represented by

$$n_{2l+1}'(s) = \frac{1}{\hat{\Lambda}(s) d_p(s)} \left[d_p(s) \,,\, \cdots \,,\, s^l d_p(s) \,,\, n_p(s) \,,\, \cdots \,,\, s^{l-1} n_p(s) \right]^T \tag{a.6}$$

and the matrix $\bar{Z}_{2l+1} \in C^{(2l+1)\times(2l+1)}$ by

$$\bar{Z}_{2l+1} := \left[n_{2l+1}(0) \,,\, n_{2l+1}'(j\omega_1) \,,\, n_{2l+1}'(-j\omega_1) \,,\, \cdots \,,\, n_{2l+1}'(j\omega_l) \,,\, n_{2l+1}'(-j\omega_l) \right] \tag{a.7}$$

By assumption (D4), for given l, there exists a set of frequencies $(0,\omega_1,\cdots,\omega_l,-\omega_1,\cdots,-\omega_l)$ such that the matrix $\bar{Z}_{2l+1}$ is nonsingular. Hence, by the first part of this proposition, we can conclude that $\bar{Z}_{2l+1}$ is nonsingular for almost all $(\omega_1,\cdots,\omega_l)^T \in R^l$. Finally, using (a.2) and (a.5), namely,

$$Z_{2l+1}' := \left[n'(0) \,,\, n'(j\omega_1) \,,\, n'(-j\omega_1) \,,\, \cdots \,,\, n'(j\omega_l) \,,\, n'(-j\omega_l) \right] \tag{a.8}$$

$$Z_{2l+1} = T^{-1} Z_{2l+1}' \tag{a.9}$$

one can easily conclude the result. This completes the proof.

Q.E.D.

Proof of Corollary 5.3:

Rewriting (5.14), we obtain the following matrix representation

$$\theta_0^T \left[n(0), n(j\omega_1), n(-j\omega_1), \cdots, n(j\omega_l), n(-j\omega_l) \right]$$

$$= \left[\hat{P}(0), \hat{P}(j\omega_1), \hat{P}(-j\omega_1), \cdots \hat{P}(j\omega_l), \hat{P}(-j\omega_l) \right] \tag{a.10}$$

or

$$\theta_0^T Z_{2l+1} = g^*_{2l+1} \tag{a.11}$$

By Proposition 5.2, Z_{2l+1} is of full column rank for almost all $(\omega_1,\cdots,\omega_l)^T \in R^l$; in which case, it can be verified that

$$\theta_0 := Z_{2l+1}(Z^*_{2l+1}Z_{2l+1})^{-1}\, g_{2l+1} \tag{a.12}$$

is a solution to the matrix equation (a.11) and, in turn, the set of equations (5.14).

Q.E.D.

Proof of Proposition 5.4:

Due to (5.35) and (5.36), the bound on the difference $\theta_T - \theta^*$ can be found to be

$$\| \theta_T - \theta^* \| = \| (R_{w^*}(0) + \Delta R_w(0))^{-1}(\Delta R_{y_p w}(0) - \Delta R_w(0)\theta^*) \|$$

$$\leq \frac{\| R_{w^*}(0)^{-1} \|}{1 - \| R_{w^*}(0)^{-1} \| \, \| \Delta R_w(0) \|} \left(\| \Delta R_{y_p w}(0) \| + \| \Delta R_w(0) \| \, \|\theta^*\| \right)$$

$$= \frac{1/\underline{\lambda}[R_{w^*}(0)]}{1 - \| \Delta R_w(0) \| / \underline{\lambda}[R_{w^*}(0)]} \left(\| \Delta R_{y_p w}(0) \| + \| \Delta R_w(0) \| \, \|\theta^*\| \right) \tag{a.13}$$

Refer to (5.7)

$$n(s) = \begin{bmatrix} (sI-\Lambda)^{-1}b \\ (sI-\Lambda)^{-1}b\hat{P}^*(s) \\ 1 \end{bmatrix} + \begin{bmatrix} 0 \\ (sI-\Lambda)^{-1}b\Delta\hat{P}(s) \\ 0 \end{bmatrix} \tag{a.14}$$

where the column matrices on the RHS are transfer functions from $r(t)$ to w^* and Δw respectively. By assumptions and (5.33), (5.34) we can estimate the following bounds:

$$\| \Delta R_w(0) \| \leq \sum_{j=-s}^{s} \left[2K_\Lambda \sqrt{1+(1+L_i^2)K_\Lambda^2} + K_\Lambda^2 l_i(\omega_j) \right] l_i(\omega_j) \frac{r_j^2}{a_j}$$

$$\leq \left[2K_\Lambda \sqrt{1+(1+L_i^2)K_\Lambda^2} + K_\Lambda^2 l_{i\infty} \right] l_i(\overline{\omega}) F_\infty$$

$$:= \alpha_1(\overline{\omega}) \tag{a.15}$$

and

$$\| \Delta R_{y_p w}(0) \| \leq \sum_{j=-s}^{s} \left[\sqrt{1+(1+L_i^2)K_\Lambda^2} + K_\Lambda l_i(\omega_j) + K_\Lambda L_i \right] l_i(\omega_j) \frac{r_j^2}{a_j}$$

$$\leq \left[\sqrt{1+(1+L_i^2)K_\Lambda^2} + K_\Lambda (l_{i\infty} + L_i) \right] l_i(\overline{\omega}) F_\infty$$

$$:= \alpha_2(\overline{\omega}) \tag{a.16}$$

As a result, the bound on $\|\theta_T - \theta^*\|$ which is defined in (a.13) can thus be simplified as follows.

$$\| \theta_T - \theta^* \| \le \frac{\alpha_1(\varpi)\|\theta^*\| + \alpha_2(\varpi)}{\underline{\lambda}[R_{w^*}(0)] - \alpha_1(\varpi)} \tag{a.17}$$

From assumptions

$$\beta_1 \le \| \theta^* \| \le \beta_2 \tag{a.18}$$

for some β_1, $\beta_2 > 0$ and

$$0 < \frac{\alpha_1(\varpi)\beta_2 + \alpha_2(\varpi)}{\beta_1(\underline{\lambda}[R_{w^*}(0)] - \alpha_1(\varpi))} \le \eta \tag{a.19}$$

for some small $\eta > 0$, the result

$$\| \theta_T - \theta^* \| \le \eta \|\theta^*\| \tag{a.20}$$

directly follows from (a.17). This completes the proof.

Q.E.D.

Proof of Proposition 5.7:

Recall that $\bar{\pi}(s)$, defined in section 4 as the transfer function from the input to the signal vector w_m, can be reformed into the following

$$\bar{\pi}(s) := \begin{bmatrix} 1 \\ \bar{v}_{p^*}(s)\hat{M}(s) \end{bmatrix} = v_p(s) + \begin{bmatrix} 0 \\ (sI - \Lambda)^{-1} b \hat{M}(s) \hat{P}^{-1}(s) \\ 0 \\ 0 \end{bmatrix} L(s)$$

$$= v_p(s) - \Delta n(s) \tag{a.21}$$

Referring to (5.51), (5.57), (5.58), we have the nominal parameter θ^* satisfy the following family of equations

$$\theta^{*T} \bar{\pi}(j\omega_i) = \hat{P}^{*-1}(j\omega_i)\hat{M}(j\omega_i)$$

From (a.21), we have

$$\theta^{*T} v_p(j\omega_i) = \theta^{*T}\, \bar{\pi}(j\omega_i) + \theta^{*T}\Delta n(j\omega_i) = \hat{P}^{*-1}(j\omega_i)\hat{M}(j\omega_i) + \theta^{*T}\Delta n(j\omega_i)$$

$$:= \hat{P}^{-1}(j\omega_i)\hat{M}(j\omega_i) + \hat{E}(j\omega_i) \qquad i = -s, \ldots, -1, 1, \ldots, s \tag{a.22}$$

where $\hat{E}(s)$ is defined as

$$\hat{E}(s) = \theta^{*T}\Delta n(s) + \hat{P}^{-1}(s)\hat{M}(s)L(s) \tag{a.23}$$

Define the vector $\Delta g_{2s} \in R^{2s}$ as follows

$$\Delta g_{2s}^T = \left[E(j\omega_1), E(j\omega_1), \cdots, E(j\omega_s), E(j\omega_{-s})\right] \tag{a.24}$$

and let $\bar{Y}_{2s}$ and ΔY_{2s} denote the matrices similar to those defined in (5.60) but corresponding to the transfer functions $\bar{n}(s)$ and $\Delta n(s)$.

Subtracting (5.58) by (a.22), we can express the bound on $\|\theta_0 - \theta^*\|$ by

$$\| \theta_0 - \theta^* \| \le \| ((\bar{Y}_{2s}+\Delta Y_{2s})W(\bar{Y}_{2s}+\Delta Y_{2s})^*)^{-1}(\bar{Y}_{2s}+\Delta Y_{2s})W\Delta g_{2s} \|$$

$$\le \frac{\| (\bar{Y}_{2s} W \bar{Y}_{2s}^*)^{-1} \|}{1 - \| (\bar{Y}_{2s} W \bar{Y}_{2s}^*)^{-1} \| \| \Delta D \|} (\|(\bar{Y}_{2s}+\Delta Y_{2s})W\Delta g_{2s}\|)$$

$$= \frac{1/\underline{\lambda}[\bar{Y}_{2s}W\bar{Y}_{2s}^*]}{1 - \| \Delta D \| / \underline{\lambda}[\bar{Y}_{2s}W\bar{Y}_{2s}^*]} (\|(\bar{Y}_{2s}+\Delta Y_{2s})W\Delta g_{2s} \|) \tag{a.25}$$

where $\Delta D = \bar{Y}_{2s}W\Delta Y_{2s}^* + \Delta Y_{2s}W\bar{Y}_{2s}^* + \Delta Y_{2s}W\Delta Y_{2s}^*$. Let $W \in R^{2s \times 2s}$ be a diagonal weighting matrix

$$W = \operatorname{diag}\left(\frac{a_1^2}{t_1}, \frac{a_{-1}^2}{t_{-1}}, \cdots, \frac{a_s^2}{t_s}, \frac{a_{-s}^2}{t_{-s}}\right) \tag{a.26}$$

where t_i is defined similarly as in (5.25) and $\frac{a_i^2}{t_i}$ represents the spectral weight of the frequency $j\omega_i$. Then, $\bar{Y}_{2s}W\bar{Y}_{2s}^*$ is simply an autocovariance matrix, i.e.

$$R_{w_m}(0) = \bar{Y}_s W \bar{Y}_s^* = \sum_{i=-s}^{s} \left[\bar{n}(j\omega_i)\bar{n}^*(j\omega_i)\right] \frac{a_i^2}{t_i} \tag{a.27}$$

In order to relate the bound in (a.25) with the frequency content of the reference input and the plant uncertainty, we first estimate several bounds as follows.

First of all, note that

$$\| (j\omega I - \Lambda)^{-1} b \| \le J_\Lambda \qquad \text{for all } \omega \in R \tag{a.28}$$

Then it is true that the following hold.

(*i*)

$$\| \bar{n}(j\omega_i) \| \le \sqrt{1 + L_c^2(1 + J_\Lambda^2 + J_\Lambda^2 l_{c2}^2(\omega_i))}$$

$$:= \alpha_3(\omega_i) \tag{a.29}$$

(*ii*)

$$\| (\bar{Y}_{2s}+\Delta Y_{2s})W\Delta g_{2s} \| \le \| \bar{Y}_{2s}W\Delta g_{2s} \| + \| \Delta Y_{2s}W\Delta g_{2s} \|$$

$$\le \left\{ \alpha_3(\bar{\omega})L_c l_{c2}(\bar{\omega})J_\Lambda + \frac{L_c^2 l_{c2}^2(\bar{\omega})J_\Lambda^2\, l_{c1}(\bar{\omega})}{1 - l_{c1}(\omega^0)} \right\} \frac{l_{c1}(\bar{\omega})F_{\infty}\|\theta^*\|}{1 - l_{c1}(\omega^0)}$$

$$:= \alpha_4(\bar{\omega})\|\theta^*\| \tag{a.30}$$

(iii)
$$\| \Delta D \| \leq 2\|\overline{Y}_{2s} W \Delta Y_{2s}\| + \|\Delta Y_{2s} W \Delta Y_{2s}\|$$

$$\leq \left[2\alpha_3(\overline{\omega}) L_c l_{c2}(\overline{\omega}) J_\Lambda + L_c^2 l_{c2}^2(\overline{\omega}) J_\Lambda^2 \right] \frac{l_{c1}(\overline{\omega}) F_\infty}{1 - l_{c1}(\omega^0)}$$

$$:= \alpha_5(\overline{\omega}) \qquad \text{(a.31)}$$

With these bounds obtained in (a.29)-(a.31), the bound on $\|\theta_0 - \theta^*\|$ in (a.25) can thus be simplified to the following

$$\| \theta_0 - \theta^* \| \leq \frac{\alpha_4(\overline{\omega}) \|\theta^*\|}{\underline{\lambda}[R_{w_m}(0)] - \alpha_5(\overline{\omega})} \qquad \text{(a.32)}$$

Consequently, given any small number $\eta > 0$, if $\overline{\omega}$ can be chosen such that

$$0 < \frac{\alpha_4(\overline{\omega})}{\underline{\lambda}[R_{w_m}(0)] - \alpha_5(\overline{\omega})} \leq \eta \qquad \text{(a.33)}$$

then the result

$$\| \theta_0 - \theta^* \| \leq \eta \, \|\theta^*\| \qquad \text{(a.34)}$$

directly follows from (a.32). This completes the proof.

Q.E.D.