INFORMATION TO USERS

This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps. Each original is also photographed in one exposure and is included in reduced form at the back of the book.

Photographs included in the original manuscript have been reproduced xerographically in this copy. Higher quality 6" x 9" black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.

UMI

A Bell & Howell Information Company
300 North Zeeb Road, Ann Arbor MI 48106-1346 USA
313/761-4700 800/521-0600

Frequency Response Based System Identification and Controller Tuning

Ali Yousef ©

September 15, 1997

A THESIS SUBMITTED IN PARTIAL FULFILMENT OF THE
REQUIREMENTS OF THE MScEng DEGREE
IN
CONTROL ENGINEERING
FACULTY OF ENGINEERING
LAKEHEAD UNIVERSITY
THUNDER BAY, ONTARIO

Abstract

A detailed simulation based analysis of different frequency response and transfer function identification techniques are performed and critically studied. Standard closed loop and open loop recursive least squares system identification techniques are studied along with methods based on Laguerre filter, frequency sampling filter and bandpass filter. Based on this critical study, the bandpass filter approach is shown to have relatively more advantages than disadvantages over the other techniques of system identification. Methods for PID controller tuning using finite and discrete set of frequency response points are developed for SISO systems and extended to MIMO systems. The controller tuning methods are supported with simulations and experiments on a temperature control system and a distillation column. Based on the results obtained, it is concluded that these methods are practical and capable of being used in industrial process control.

Acknowledgements

I dedicate this work to my great mentors Drs. K. Natarajan and A.F. Gilbert for their help and support, and to my parents who always stood by me, and to my good friends H. Golestani and A. Akoum.

The financial support of the network of centers of excellence (NCE) in mechanical and chemimechanical wood pulps to my supervisors and their financial support of me during the course of this work made it possible.

Contents

List of Figures

Chapter 1

A Comparative Study of Some System Identification Methods

1.1 Introduction

System identification is the obtaining of parsimonious quantitative data about a process which can be used to predict closed loop (or open loop) behaviour due to setpoint change, disturbance and noise. Almost all methods of identification involve changing some input and measuring the output response. Due to the presence of disturbances and noise, repetitive changes to the input are often always required to separate the process response from disturbance and noise effects. Three configurations of system identification are possible: closed loop, open loop and true open loop. In closed loop identification changes from the setpoint to the closed loop process output are used to identify the closed loop system from which the known controller dynamics are factored out to obtain process information. In open loop identification, with the controller in automatic mode, the changes in the manipulated variable and process output are used to identify the process dynamics. In this case the loop is still closed and setpoint changes are applied to effect changes in manipulated variable and process response. In true open loop identification, the controller is put on manual and changes are applied directly to the manipulated variable and process response is used to identify the dynamics of the process.

From the point of view of process regulation at an operating point, the process is identified assuming linear time invariance. However in such cases too, small and symmetric repetitive changes of the input are needed to separate the effect of disturbances and noise and to average the process response on the up and down bumps of the manipulated variable. For example, in a temperature control system, active heating behaviour is different from passive cooling behaviour even in a small signal model. In this chapter as in most previous work, it is the small signal process behaviour that is identified. Classically this is carried out by using parametric transfer function model such as gain + first order time constant + dead-time or in a non-parametric form such as Bode plots to which subsequently a transfer function is fitted.

The parametric form often requires the knowledge of at least the order of the process and is usually faster to identify. Process identification can be carried out in the true open loop form, open loop form or closed loop form. However, in any closed loop (continuous time) scheme, since dead-time enters into the numerator and denominator of the Laplace domain transfer function separating time constant(s) from deadtime is not easily accomplished if they are close to each other. If a closed loop scheme with process transfer function in the sampled data domain (z-domain) (of known order and unknown coefficients) is used with recursive least squares (RLS) [1] to identify the coefficients, it is often the case that the regression vector has low signal to noise ratio and the estimation procedure is unreliable. The regression vector in this case often involves the difference between successive process outputs/inputs which causes the low signal to noise ratio. For closed loop identification of the process, however, one can assume a closed loop transfer function with unknown coefficients in the z-domain and identify its coefficients directly using recursive least squares. Following this stage of identification, using the known controller transfer function, the loop can be opened to obtain the transfer function of the process.

In the non-parametric form, closed loop schemes are easier to use as the closed loop (continuous-time) frequency response can be obtained and the controller frequency response used to calculate the process frequency response. The traditional methods of applying sinusoids of different known frequencies at the input and measuring the output sinusoid amplitude and phase shift to obtain Bode plots is a well known technique in this category. While this approach can be speeded up by using Fourier analysis (fast Fourier transforms [2, 8]), the non-parametric approach is often slower than parametric system identification.

True open loop identification as well as non-parametric approaches such as frequency response schemes are often batch type identification schemes rather than on-line recursive schemes. The main disadvantage of batch schemes is often that one finds out the identification experiment did not reveal the information sought until after the experiment is terminated. If the lack of information is detected while the experiment is in progress, it would enable corrective efforts to be applied immediately (such as amplitude of the bumps/ frequency of the bumps etc.). Also, from the point of the use of the identification results in adaptive control, fault-diagnosis and intelligent control, it is often desired to have on-line recursive identification schemes. In this chapter therefore only the recursive system identification based on open loop or closed loop schemes is considered.

As mentioned earlier, in their most commonly used form, recursive identification procedures involve a parametric transfer function in the z-domain of known order whose parameters are estimated recursively using least squares [1]. The order of the transfer function has to be known apriori requiring preliminary identification. While combined parameter and order estimation are possible [3], these recursive algorithms are necessarily high dimensional. It is almost always impossible to relate the z-domain transfer function to the continuous-time transfer function of the process. The z-domain identification can be carried out in closed loop or open loop fashion. Since this method of identification is

the standard against which all other schemes are compared, in the next section we consider the performance of recursive least squares estimation of z-domain transfer functions. It is argued there that closed loop schemes are preferable to open loop schemes in this case.

A fundamental argument advocating careful use of this approach is that in the z-domain approach it is usually assumed that the controller is a sampled-data controller with a zero-order hold at its output and that the identification experiment is carried out at the same rate in synchrony with the control loop. In industrial process control, the controller is usually implemented in a distributed control system (DCS) where control calculations take place at a high sampling rate. The controller for each loop appears to be operating in a continuous-time manner in relation to process dynamics. The facility to add real-time identification code into a particular controller is often limited. Data logged into a database by a central (higher) DCS processor is usually at a slower rate than the control rate. If this logged data is used for identification then the zero-order hold assumption of the identification experiment is not valid. If the identification is carried out on a separate (portable) computer different from the control computer, synchronization to the control computer is next to impossible and again the identification of discrete-time transfer functions as set out [1] is not possible. If facility was available to add the identification code into the control processor, then due to the high sampling rate of this processor and the variable dead-time of processes, the order of the parameterized transfer function is high leading to a high dimensional and time consuming least squares. Some form of averaging of inputs and outputs is necessary to reduce the order of the transfer function. If a closed loop scheme of identification is used in this manner, then it is not directly possible to open the loop to compute the process transfer function due to the implicit presence of two rates - controller update and RLS update. Open loop schemes are therefore preferable in this case. However in this case as discussed in the next section the RLS algorithms have to be used carefully.

The organization of this chapter is as follows: In the next section, the standard RLS algorithm for process identification of transfer functions is discussed by simulations. In the three sections following three relatively new identification methods, Laguerre filter method [7, 9, 6, 12], frequency sampling filter method [5, 4, 10, 11] and bandpass filter methods [15] are studied through the same simulations. Section 7 concludes the chapter.

1.2 Transfer Function Estimation

Standard RLS schemes parameterize the transfer function in the z-domain. To illustrate some of the various possibilities in the use of RLS schemes a process is simulated. The simulated process consists of a continuous time process with transfer function given in Equation (1.1) along with a PI controller (Equation (1.2)). The PI controller calculates manipulated variable updates every 50 ms. The manipulated variable is held at its last value until the next controller update (zero-order hold). The PI controller transfer function in the z-domain is given in Equation (1.2) where T denotes the sampling period and

K_P and K_I are the PI gains. The process parameter variations and their nominal values are assumed to be given by Equation (1.3). The nominal value of the PI gains are also given in Equation (1.3).

$$G(s) = K\frac{e^{-s\tau_d}}{1+s\tau} \tag{1.1}$$

$$C(z) = K_P + \frac{K_I T}{1-z^{-1}} \tag{1.2}$$

$$\begin{gathered} 0.5 \leq K \leq 2, \;\; 0.4 \leq \tau_d \leq 0.8, \;\; 0.5 \leq \tau \leq 2 \\ K_{nom} = 1, \;\; \tau_{dnom} = 0.6, \;\; \tau_{nom} = 1 \\ K_P = 0.2 \;\; K_I = 0.5 \end{gathered} \tag{1.3}$$

Load disturbances are assumed to act on the process output additively and have a dynamics given by

$$d(t) = A_{dis} e^{-t/\tau_{dis}} \cos(\omega_{dis} t) \tag{1.4}$$

In Equation (1.4) A_{dis}, τ_{dis} and ω_{dis} are the amplitude of the disturbance, the decay rate of the disturbance and the frequency of the disturbance in rad/s respectively. Time t in Equation (1.4) is measured from the start of the disturbance (and not from the start of the simulation). A cosine form is used as setting the ω_{dis} to zero, it is possible to generate exponentials and by setting large value for τ_{dis} relative to total simulation time, one can generate step disturbances.

The process is simulated in a digital computer with a fine time step of 1 ms and all time constant/time delay (changes are made to) have a minimum resolution of the same 1 ms. For both open loop and closed loop schemes, the setpoint is excited by a square waveform with amplitude ± 1 and a period of 30 s. The period selected ensures that the closed loop system comprised of the process operating with nominal parameters and the PI controller tracks the setpoint before the next bump is applied *i.e.* steady state is achieved. Simulations are run for a duration of 5 periods of the excitation. To study the effects of disturbance, it is applied halfway through the simulation. Unless otherwise specified disturbance is a decaying exponential with amplitude of 1 and decay time constant of 400 s in Equation (1.4). Process output measurements are assumed to be corrupted by Gaussian random noise with standard deviation of 0.025. Setpoint waveforms have no noise added. Unless otherwise specified all simulations reported in this chapter are carried out in this manner.

To illustrate the accuracy of RLS when model is matched to the process, a matched z-domain transfer function obtained by discretizing the first order + dead-time process with nominal parameters is considered in Equation (1.5)

$$G(z) = \frac{\mu_0}{z^{13}(1-\alpha z^{-1})} \tag{1.5}$$

The true values of μ_0 and α are 0.0488 and 0.9512 respectively. When RLS is run in an open loop scheme with initial diagonal covariance of 100 and forgetting factor of 0.999

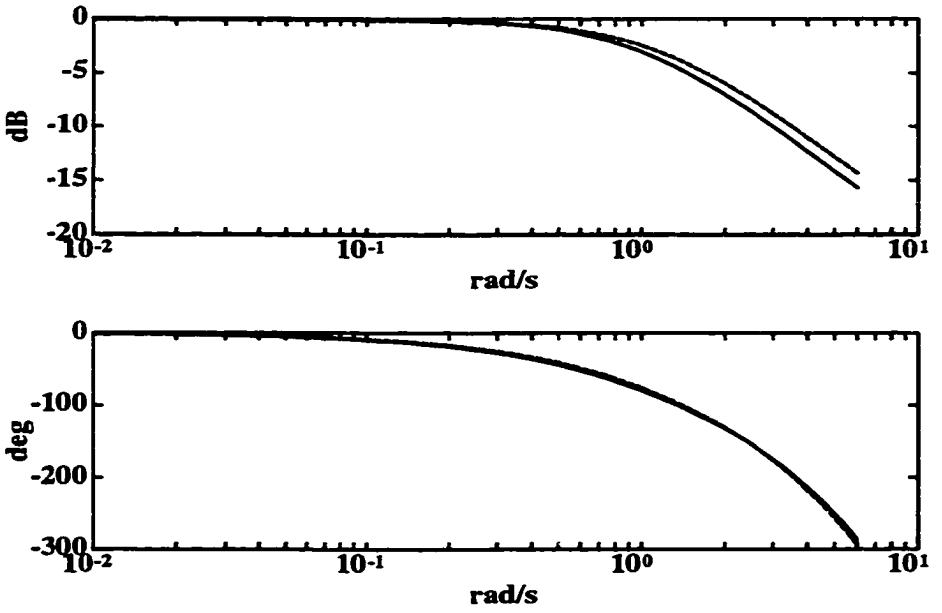


Figure 1.1: Frequency responses of nominal process and matched discrete-time transfer function identified by open loop RLS scheme.

after 3000 samples (5 periods of excitation) at a sampling period of 50 ms, the estimated values of μ_0 and α are 0.05693 and 0.9424 respectively. A plot of the frequency response of the discrete transfer function $G(z)$ and the the continuous time process transfer function up to about 6 rad/s is shown in Figure 1.1. The solid line in Figure 1.1 is the true continuous time process frequency response while the dashed line is the (discrete-time) frequency response computed from the estimates. Note that the frequency of 6 rad/s is approximately 0.05 (1/20) of the sampling frequency (20Hz) in rad/s. It indicates reasonably good match between the discrete-time and continous time frequency responses. Next, a simulation is carried out to include the effect of load disturbance on the above estimates. The estimates from RLS after one period (of the setpoint waveform) after the load disturbance hits the process are now $\mu_0 = 0.0097$ and $\alpha = 0.9901$ respectively. Now the estimated frequency response is off in the same frequency range by a significant margin as shown in Figure 1.2. Once again the solid line in Figure 1.2 is the true continuous-time process frequency response and the dashed line represents the (discrete-time) response computed from the estimates. The poor performance is due to the development of a significant bias in the manipulated variable due to the load disturbance. Asymptotically as the disturbance vanishes this frequency response matches the true response as before. However if the biased estimates are used in (adaptive) controller design then robust performance cannot be achieved in this case.

Next, consider the process operating at the lowest set of parameters $K = 0.5, \tau_d =$

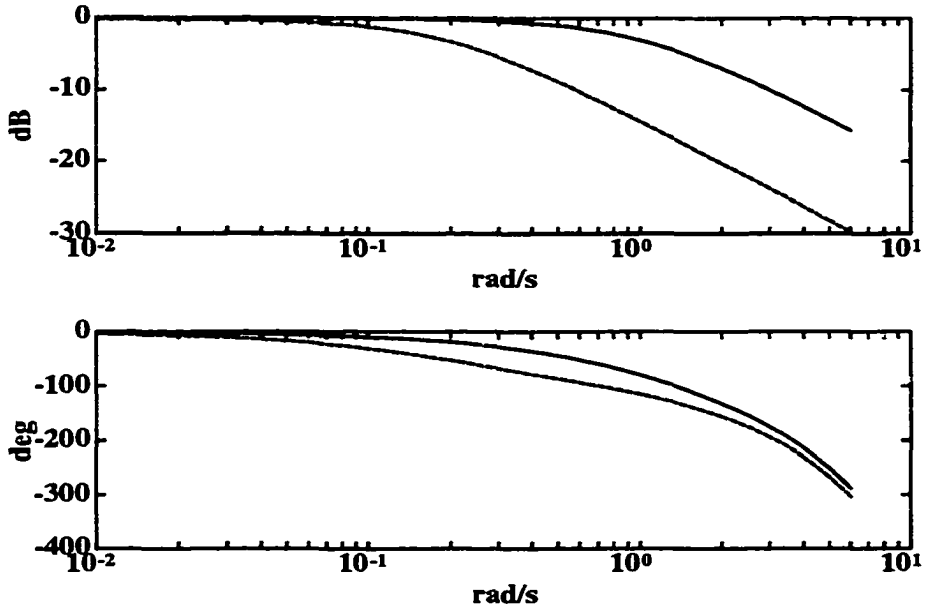


Figure 1.2: Frequency responses of nominal process and matched discrete-time transfer function identified by open loop RLS scheme in the presence of disturbance.

$0.4, \tau = 0.5$. The matched model for this case is given by Equation (1.6).

$$G(z) = \frac{\mu_0}{z^9 (1 - \alpha z^{-1})} \tag{1.6}$$

With this process model, the PI controller as before and with no load disturbance, the open loop RLS scheme after 3000 samples estimates were $\mu_0 = 0.0785$ and $\alpha = 0.8419$. The (discrete-time) frequency response of the estimated transfer function (dashed line) and the true continuous time transfer function (solid line) is shown in Figure 1.3. One immediately sees the poor fit of the estimated frequency response to the actual frequency response of the continuous time process in this case. Consider now the same identification experiment with a controller with an I gain of 2 in Equation (1.2) instead of the nominal value of 0.5. This leads to estimated values of $\mu_0 = 0.0537$ and $\alpha = 0.8921$. The values of μ_0 and α in the latter case are closer to the ideal values of 0.0476 and 0.9048 respectively than the former values. The frequency response fit for the latter case is shown in Figure 1.4 and the improvement of the fit is visible.

For the case of the process operating at its high parameter values ($K = 2, \tau_d = 0.8, \tau = 2$) and the PI controller with nominal values of gains as per Equation (1.3), the controller is very aggressive and the output response and process input is highly oscillatory. A good match is obtained between the frequency response of the estimated transfer function and the actual continuous time frequency response of the process up to 6 rad/s.

From these simulations it is seen that the more aggressive PI controller supplies a richer excitation to the process enabling better identification in the higher frequencies than a less aggressive controller. However the use of such an aggressive controller would not

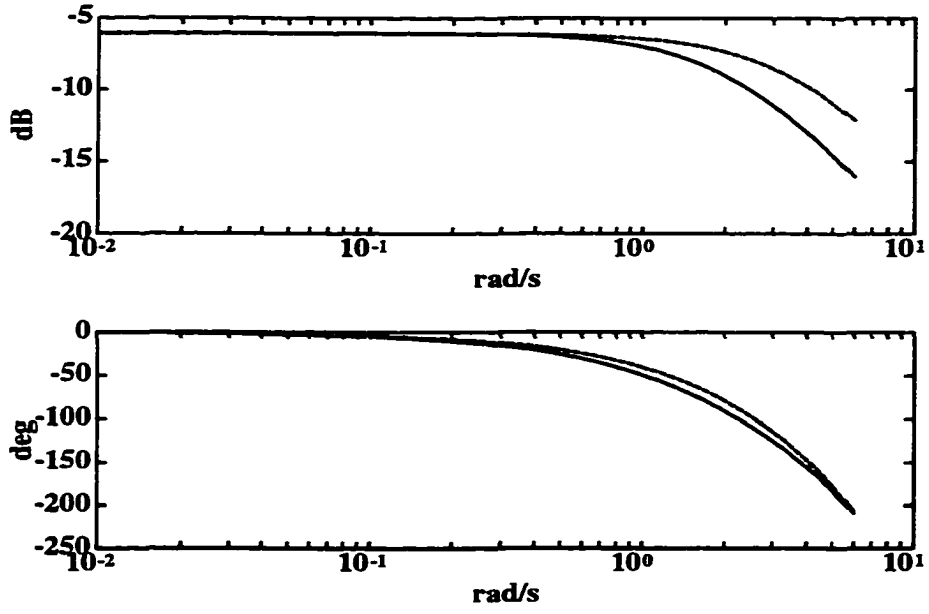


Figure 1.3: Frequency responses of process with parameters $K = 0.5, \tau_d = 0.4, \tau = 0.5$ and matched discrete-time transfer function identified by open loop RLS scheme with PI gains of 0.2 and 0.5.

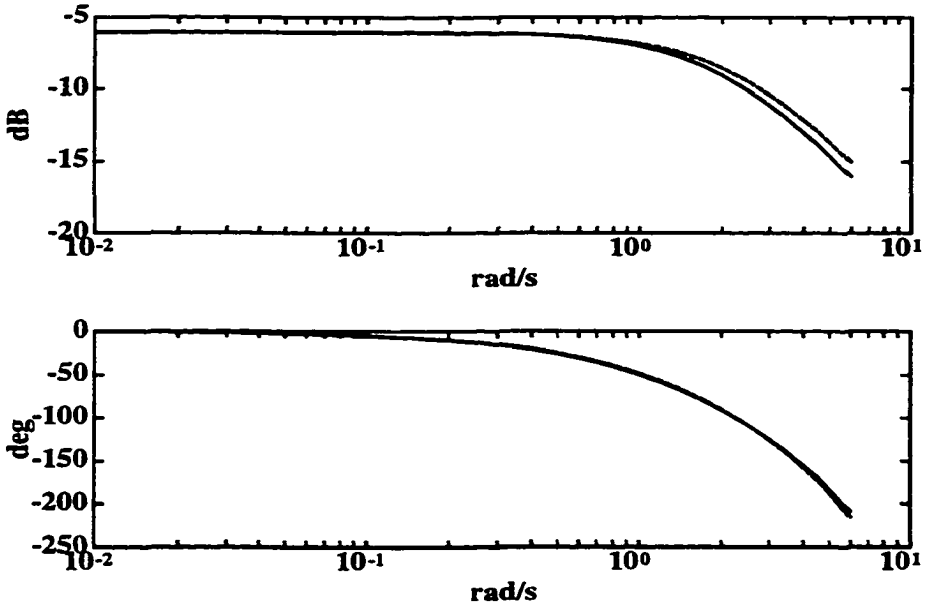


Figure 1.4: Frequency responses of process with parameters $K = 0.5, \tau_d = 0.4, \tau = 0.5$ and matched discrete-time transfer function identified by open loop RLS scheme with PI gains of 0.2 and 2.

be practical if the process exhibits significant and sudden gain and dead time variations. Finally, determining apriori, the correct controller gains in a practical identification experiment is an issue. This is often the purpose of determining the transfer function. Thus from a practical stand point open loop identification has to be used carefully when the process is controlled with PI controllers. Lack of richness of excitation of the process due to a (detuned) controller or the presence of load disturbances can produced significantly biased results.

Assuming that the PI controller is satisfactory, a possible way out of the problem of load disturbance, suggested in the standard RLS literature, is the use of transfer functions between the backward difference in manipulated variable to backward difference in process output (velocity mode RLS). This may help in processes where the bumps are applied before the process settles. When the bump period is large and process input and output settle before the next bump, then backward differencing may provide no excitation to the least squares for considerable periods of time. This lack of excitation causes poor parameter estimation and is not a good solution to the problem of disturbance.

An attempt using the backward difference approach in RLS for the simulation example with process and PI controller operating with nominal parameters (Equation (1.3) and process model transfer function as per Equation (1.5) leads to estimates of $\mu_0 = 0.3306$ and $\alpha = -0.3575$ respectively (after 3000 samples) which are considerably off from their true values.

The main conclusion from the previous simulation experiments is that open loop RLS estimation of conventional transfer functions of processes controlled by PI (PID) controllers, even in the matched transfer function case is fraught with the danger of lack of persistency and richness of of the excitation at the input to the process.

An alternative is to perform closed loop identification with a parameterized transfer function and then open the loop using the controller transfer function. If $T(z), C(z)$ denote the closed loop transfer function and controller transfer function in a unity feedback system, then the open loop transfer function is given by:

$$G(z) = \frac{T(z)}{1 - T(z)} \frac{1}{C(z)} \tag{1.7}$$

Turning now to closed loop RLS identification between setpoint and process output, the first issue is parameterization of the transfer function. Two approaches are possible. In the first assuming the parameterization of the the open loop process transfer function, compute the structure of closed loop transfer function. For the nominal process model represented by Equation (1.5) and the nominal PI controller, the closed loop transfer function is:

$$T(z) = \mu_0 \frac{(K_P + K_I T) z^{-d} - K_p z^{-(d+1)}}{[(1 - z^{-1}) - \alpha z^{-1} (1 - z^{-1}) + \mu_0 \{(K_P + K_I T) z^{-d} - K_p z^{-(d+1)}\}]} \tag{1.8}$$

It is possible to set up a RLS formulation from Equation (1.8) to estimate the coefficients μ_0, and α. However as Equation (1.8) reveals, the regression vector involves (scaled) differencing operations and usually, therefore, has low signal to noise ratio causing unreliable

estimation of coefficients. This approach is therefore not very useful. A second approach to parameterization of the closed loop transfer function is to over parameterize it as:

$$T(z) = \frac{b_0 z^{-d} + b_1 z^{-(d+1)}}{1 + \sum_{i=1}^{d+1} a_i z^{-i}} \tag{1.9}$$

The choice of two coefficents in the numerator is motivated by the PI controller and Equation (1.8). The denominator however is considerably over parameterized. It is no longer possible to relate the estimated coefficients to the process parameters. However one expects that the invariants of impulse response and equivalently of the frequency response of the closed loop system will be well captured by Equation (1.9). To illustrate this, a simulation of the nominal process and PI controller with no load disturbance and $d = 13$ and 7 in Equation (1.9) are carried out. The closed loop frequency response obtained from the estimates after 3000 samples (at a sampling period of 50 ms) is illustrated in Figure 1.5 along with the closed loop continuous-time frequency response assuming a continuous-time PI controller operating with the same gains as the sampled-data PI controller of the simulation. In Figure 1.5 the solid line is the true continuous-time closed loop response while the dashed line corresponds to the $d = 13$ case and the dotted line corresponds to the $d = 7$ case respectively. In Figure 1.6 is shown the effect of disturbance applied half way through the simulation on the frequency response of the estimated closed loop transfer function 600 samples (one period of excitation at a sampling period of 50 ms) after the disturbance hits the process for the two choices of d.

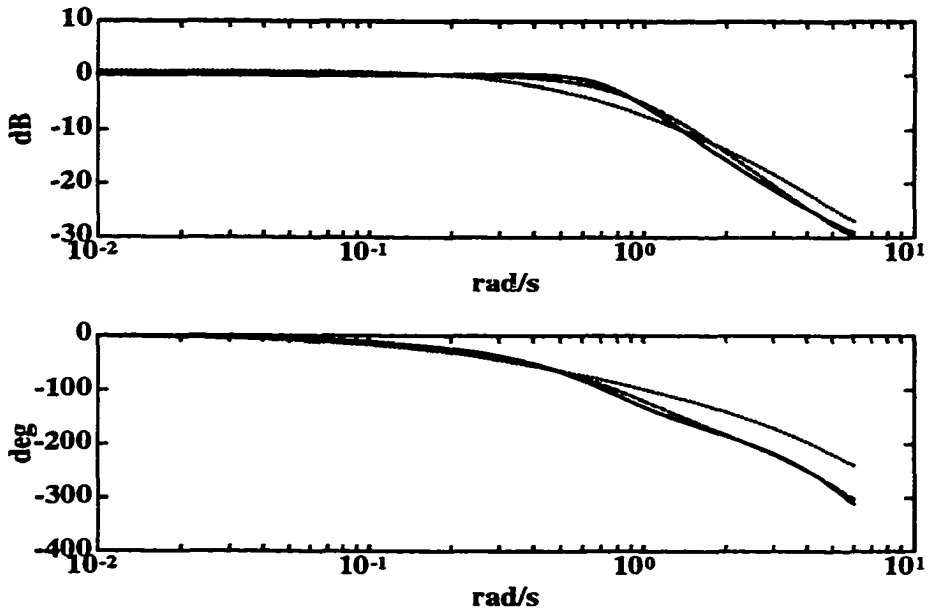


Figure 1.5: Frequency responses of nominal process and discrete-time closed loop transfer function identified by closed loop RLS scheme with $d = 13$ and $d = 7$ in Equation (1.9).

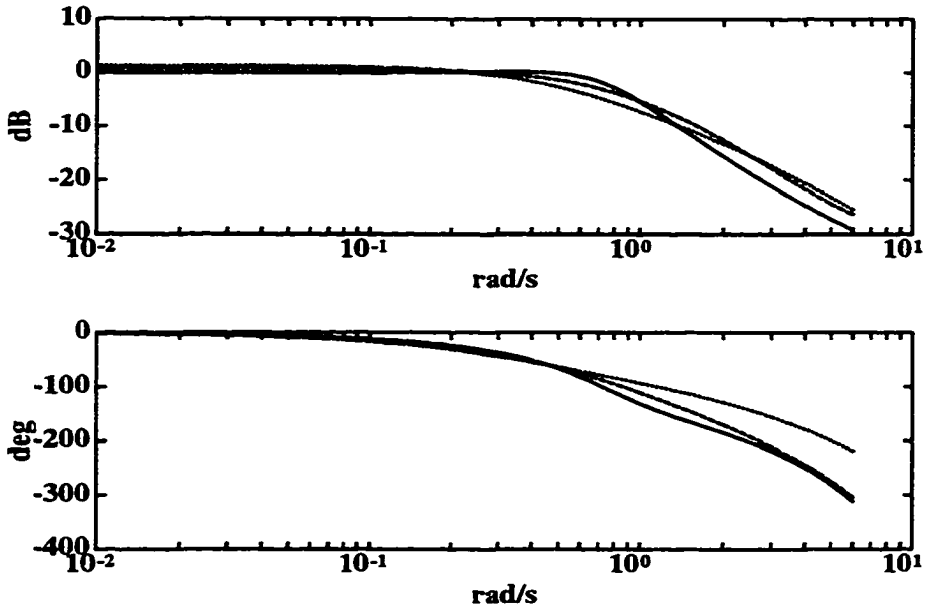


Figure 1.6: Frequency responses of nominal process and discrete-time closed loop transfer function identified by closed loop RLS scheme with $d = 13$ and $d = 7$ in Equation (1.9) in the presence of disturbance.

While Figures 1.5, 1.6 indicate the improved results obtainable by closed loop identification visavis open loop identification, to make a fair comparison one should compute the frequency response of the process using Equation (1.7) on the $d = 13$ results of these two figures. In Figure 1.7, the frequency response of the process obtained in this manner in comparison with the true process (continuous-time) frequency response is shown. One immediately sees a problem in the low frequencies.

Since the period of the excitation is 30s at the setpoint, the lowest frequency at which excitation is available to the RLS is approximately 0.2rad/s. Therefore at frequencies lower than about 0.1 rad/s, the estimated model is suspect. To confirm this consider the excitation of the setpoint by a square wave with period of 200s with no load disturbance. The computed frequency response of the process from the closed loop identification results at the end of the simulation (5 cycles of setpoint excitation) are shown in Figure 1.8 along with the true process frequency response.

The improvements in Figure 1.8 of the computed process low frequency response is clearly seen over Figure 1.7 as the lowest excitation (fundamental) frequency is now approximately 0.03rad/s. It is also seen that high frequency errors are higher in Figure 1.8 than Figure 1.7. This too is to be expected as the harmonic amplitudes of the excitation drop as the inverse of the harmonic frequencies. Hence signal to noise ratio (SNR) is lower at higher frequencies than that in Figure 1.7 where the fundamental of the excitation is at 0.2 rad/s.

While the explanation offered above is based on excitation analysis, another explana-

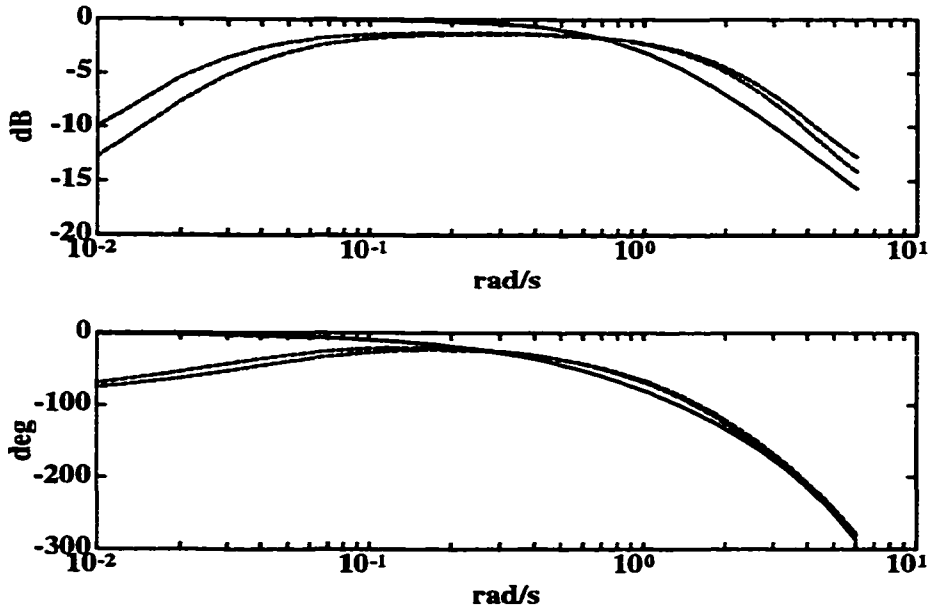


Figure 1.7: Frequency response of discrete-time process transfer function ($G(z)$) computed from closed loop identification results of Figures 1.5 and 1.6 for $d = 13$ model.

tion is that the inverted controller transfer function contributes a $(z - 1)$ factor to the numerator of the computed $G(z)$. In the absence of exact cancellation between denominator and this factor in the numerator, the low frequency response of $G(z)$ asymptotically tends to 0 as the frequency tends to zero. This explanation however does not show that the computed $G(z)$ will improve with frequency content of the excitation.

Finally the performance of a single parameterization such as Equation (1.9) with $d = 13$ for the two limit sets of process parameter values and the nominal process are shown in Figure 1.9. The good fit of the over parameterized model of Equation (1.9) to process parameter variations is clearly evident.

Several conclusions can be drawn out of these simulation experiments about the use of RLS for system identification:

1. closed loop identification with over parameterized transfer functions offers better identification of the frequency response than open loop identification. This is because this method is not sensitive to the bias due to load disturbances that develops at the process input (manipulated variable) and due to the richer excitation at the setpoint than at the process input.

2. The fundamental frequency and a frequency up to which SNR in the excitation is good are the range of frequencies over which the closed loop identification is reasonably accurate, particularly when opening the loop to compute process frequency response. This range is essentially imposed by the excitation period and the SNR of the (measured output) signal spectrum visavis inherent noise.

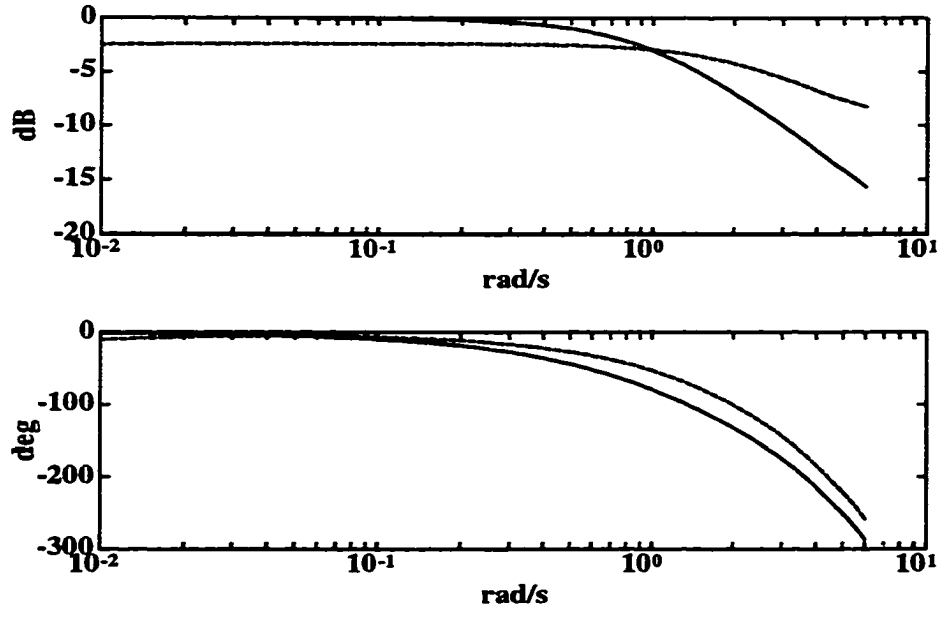


Figure 1.8: Frequency response of process transfer function ($G(z)$) computed from closed loop identification results for $d = 13$ model with setpoint excitation of period = 200.

3. The over parameterized transfer function coefficients fluctuate considerably over a simulation run. However the frequency response (equivalently impulse response, step response etc..) computed from their values at a particular time instant are quite invariant and therefore capture information about the system. The fluctuations in the coefficients of the estimation do render the computation of controller parameters for adaptive control from the coefficient values susceptible to fluctuations and therefore may not be suitable. Using the frequency response estimate directly in design along the lines of [15, 16, 13] may be a preferred alternative.

1.3 Closed loop identification with Laguerre functions

The basis of identification of transfer functions using Laguerre or other series approach such as series in z^{-1} lies in Runge's theorem in complex analysis [17] which assures convergence to the true frequency response provided increased number of terms are used in such series. In the Laguerre approach [7, 9, 12, 14] the transfer function of the closed loop systems is approximated by

$$T(s) \approx \sum_{i=1}^{i=N} c_i L_i(s) \tag{1.10}$$

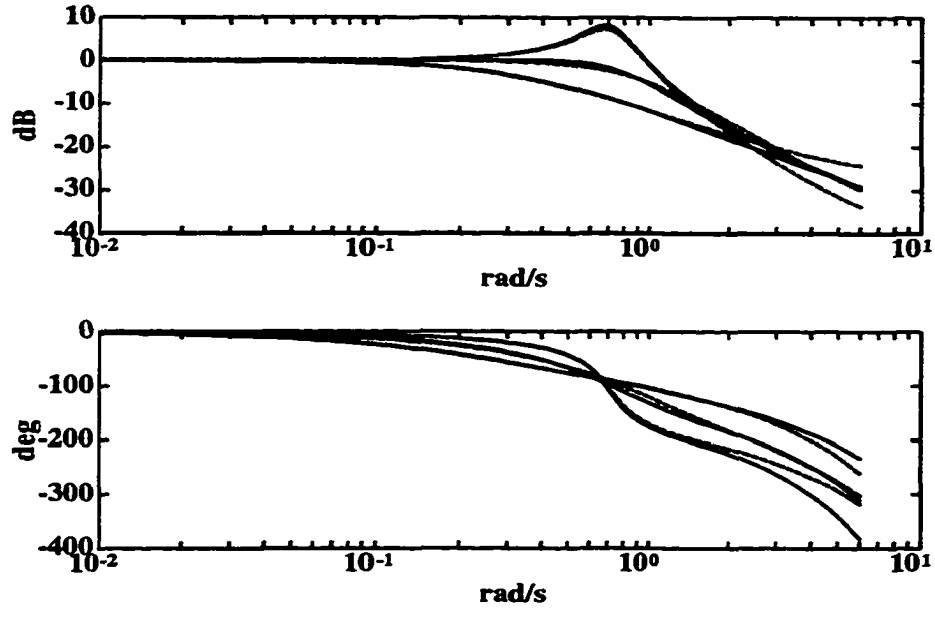


Figure 1.9: Frequency response of closed loop identification of model of Equation (1.9) with $d = 13$ for the nominal process parameter value set and the two limit sets.

where $L_i(s)$ is the Laguerre function of order i and is given by:

$$L_i(s) = \sqrt{2p}\frac{(s-p)^{i-1}}{(s+p)^i} \tag{1.11}$$

The main issues are deciding the order N and the Laguerre pole p. While Runge's theorem [17] guarantees that any $p > 0$ with a suitable order N will suffice to approximate $T(s)$, poor choice of p can lead to high N increasing on line recursive computational complexity. Several papers [9, 6] dealing with the optimum choice of p for known process transfer functions have been developed in the literature to minimize N. While they provide guidelines to the choice of p, these references often require apriori knowledge about the process.

For a given choice of p and order N, the Laguerre function model of the closed loop system can be expressed in a state space form with r as the setpoint waveform and y as the closed loop process output measurement as [7, 14]:

$$\begin{aligned} \dot{x} &= Ax + Br \\ y &= Cx = \begin{bmatrix} c_1 & c_2 & \cdots & c_i & \cdots & c_N \end{bmatrix} x \end{aligned} \tag{1.12}$$

where ($'$ denotes transpose)

$$A = \begin{bmatrix} -p & 0 & \cdots & 0 \\ -2p & -p & \cdots & \vdots \\ \vdots & \vdots & \ddots & 0 \\ -2p & -2p & \cdots & -p \end{bmatrix}_{N \times N}$$

$$B' = \sqrt{2p}\left[\begin{array}{ccccc} 1 & 1 & \cdots & 1 & 1 \end{array}\right]_{1\times N} \tag{1.13}$$

Since the setpoint is sampled once every T seconds and the standard assumption of zero order hold is made on setpoint excitation for intersampling behaviour, the above Equations can be discretized to form the sampled data description of the system as:

$$x_{k+1} = Fx_k + Gr_k \tag{1.14}$$
$$y_k = Cx_k \tag{1.15}$$

with $F = e^{AT}, G = A^{-1}[F - I]B$.

Using setpoint excitation measurements and an initial condition $x_0 = 0$ the successive states of the Laguerre network can be computed from Equation (1.12). Using RLS [1], process output measurement y_k, the Laguerre coefficients c_i can be computed recursively using the states of the Laguerre network as the regression vector.

In the simulations the closed loop process considered was the same as in the previous section of this chapter (Equations (1.1, 1.2, 1.3)). For the RLS algorithm, the initial covariance of the coefficients was set as before to 100 and a forgetting factor of 0.999 was employed as before. For the nominal process and controller parameters with no noise on setpoint measurements but a Gaussian noise level of standard deviation 0.025 on the process output measurements and no disturbance, the closed loop frequency response of the true $T(s)$ and the identified Laguerre approximation are presented in Figure 1.10 for three choices of p (0.8, 2, 5) for a given $N = 10$. These results were computed from the estimates obtained after processing 3000 samples (5 periods of excitation) at a sampling period of 50 ms as in the previous Section. In Figure 1.10, the true closed loop response is the solid line. The Laguerre approximations are shown by dashed lines. The dashed line with squiggles at the low frequency end is the $p = 0.8$ approximation. The dashed line with a dip at about 5 rad/s (in the magnitude plot) is the $p = 5$ Laguerre approximation and the third dashed line is the $p = 2$ Laguerre approximation. In Figure 1.11, the frequency response of the process computed using the closed loop Laguerre results of Figure 1.11 are shown. The description of the dashed lines for the Laguerre approximations is as given above for Figure 1.10 From these Figures, it is seen that with $p = 2$, the Laguerre coefficient estimates provide the best fit to the closed loop and open loop frequency response when compared with the other two values of p. The choice of p is critical in this method of identification. Note that unlike standard RLS transfer function estimation, for the correct choice of p the open loop process gain plots extend well below the frequency of excitation in this method. However phase response at low frequencies shows some errors even for the best choice of p.

In Figure 1.12 are the Laguerre closed loop identification results (after 3000 samples) for the choice of $p = 2$ and three choices of N (8, 10, 12). Figure 1.13 shows the results of computing the process frequency response from the Laguerre identification results of Figure 1.12. In Figures 1.12 and 1.13 the dashed lines represent the Laguerre approximation while the solid line respresents the true frequency response. The $N = 8$ approximation has a dip at 4 rad/s in the magnitude plot while the $N = 12$ approximation has squiggles

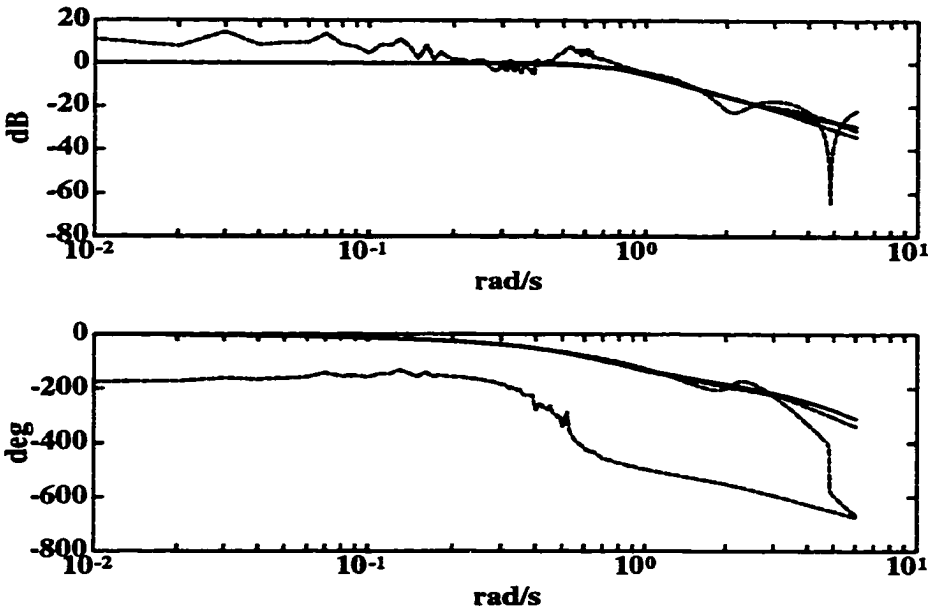


Figure 1.10: Frequency response of closed loop identification of Laguerre model with $p = 0.8, 2, 5$ and $N = 10$.

in its low frequency response. From these Figures, it is seen that once the correct choice of p is made, then the choice of N is critical too in that increasing it to a large number causes measurement noise to affect the frequency response. Too low a value causes high frequency errors and something in between provides reasonably good estimation.

In Figure 1.14 is shown the closed loop estimation for all 3 process parameter sets with the PI controller with nominal gains (Equation (1.3)) and the performance of the Laguerre estimator for the "best" choice of $p = 2$ and $N = 10$ for the process identification. The solid lines represent the true response while the dashed lines are the Laguerre approximation. For the high gain, high delay and high time constant this choice of Laguerre parameters is obviously inadequate. A complex Lagurre pole choice would be needed as the closed loop system is considerably under damped. For the low gain, low delay and low time constant process, high frequency squiggles are seen in the frequency response estimated from the Laguerre approximation. However, this frequency response estimate would probably be adequate in most controller designs. For wide variations of process parameters, it becomes necessary therefore to adjust the Laguerre pole and/or order to better fit the frequency response. The effect of disturbances in the Laguerre closed loop estimation is next studied. The estimated frequency response after 600 samples (one period of setpoint excitation) after the disturbance enters the process is shown in Figure 1.15. The solid line represents the true closed loop frequency response of the nominal process with PI controller with nominal gains (Equation (1.3)) while the dashed line represents the Laguerre approximation. The performance in this case is noticably poorer than the case of closed loop transfer function estimation using over parameterized

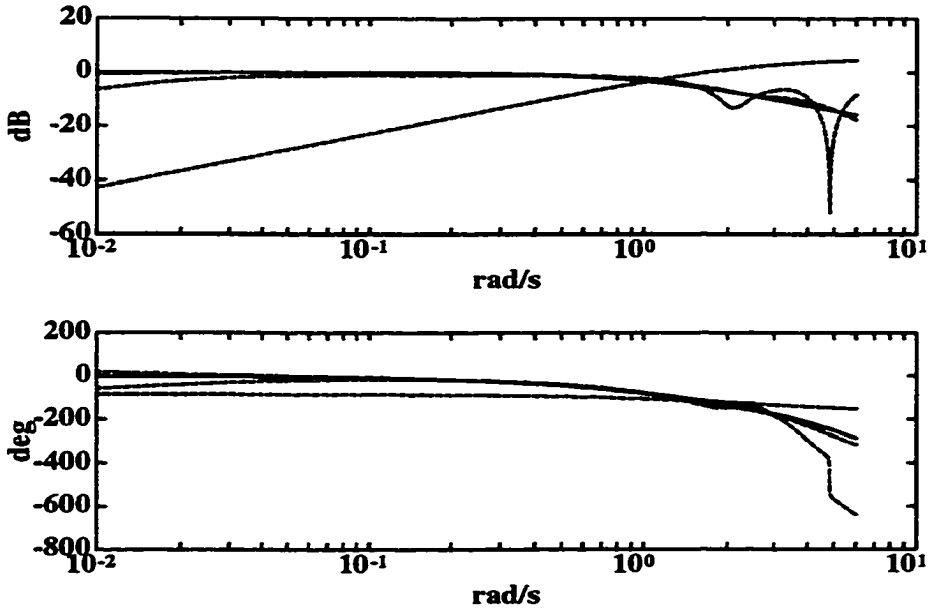


Figure 1.11: Frequency response of process computed from Laguerre identification of closed loop with $p = 0.8, 2, 5$ and $N = 10$.

transfer functions as shown in Figure 1.6.

From these simulations, the following conclusions can be drawn:

1. for the right choice of p and N (2 and 10 for simulation with the nominal process), the Laguerre approach fits the closed loop transfer function frequency response well. The open loop process frequency response computed from the closed loop measurements fit the process well in this case.

2. For increased N, noise transmission into the estimates affects the frequency response as is seen in Figures 1.12 and 1.13. However lower N, while degrading higher frequency performance does not radically change the frequency response. Hence the minimum value of N satisfactory to the identification should be used and over parameterization should be avoided.

3. The choice of p is quite important and estimates using [9, 6] should be used when adequate data is available for their use particularly in the presence of process parameter variations.

4. Disturbance effects should be minimized by continuing the identification experiment long enough before using the coefficients so that their effects are not felt.

5. Good filtering of the noise in the process measurements and on setpoint excitation (to equalize delay and gain variations introduced by the filters) within the bandwidth of the estimation sought is necessary.

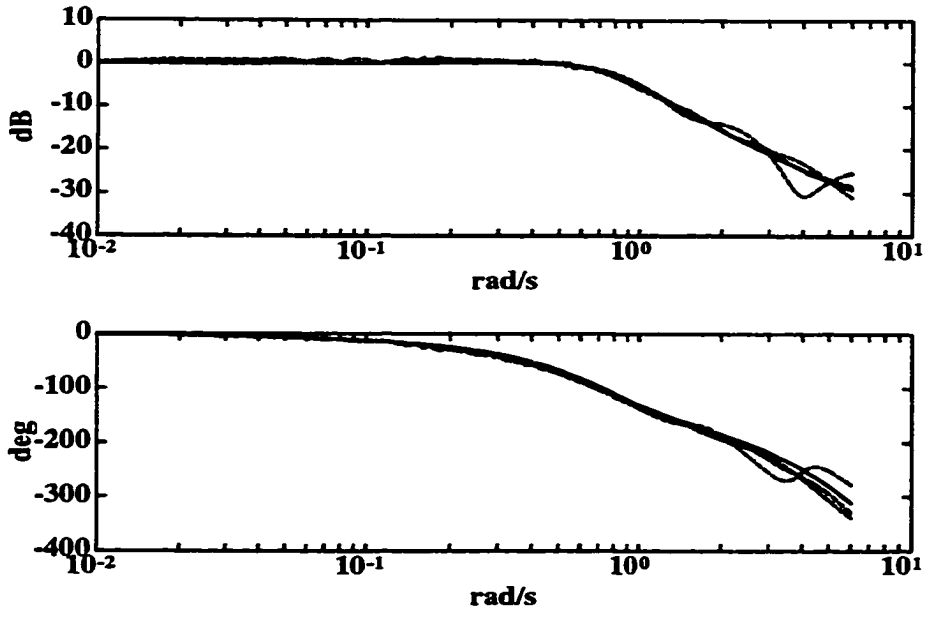


Figure 1.12: Frequency response of closed loop identification of Laguerre model with $p = 2$ and $N = 8, 10, 12$.

1.4 Closed Loop Identification with Frequency Sampling Filter

The frequency sampling filter [2] (FSF) can be used as an estimator of frequency response [4, 5, 10, 11]. In this method the discrete-time closed loop transfer function $T(z)$ is written as

$$\begin{aligned} T(z) &= \sum_{k=0}^{N-1} T\left(e^{\frac{j2\pi k}{N}}\right) H_k(z) \\ H_k(z) &= \frac{1}{N} \frac{1 - z^{-N}}{1 - e^{\frac{j2\pi k}{N}} z^{-1}} \end{aligned} \tag{1.16}$$

where $T\left(e^{\frac{j2\pi k}{N}}\right)$ is the frequency response of the transfer function at the k (normalized) frequency. The idea essentially has the notion that a process with finite impulse response of order N can be described by its frequency response values at N points and interpolation between the points with standard frequency domain functions. It consists of N parallel narrow band-pass filters, each with centre frequency at one of N equally spaced points on the z-plane unit circle, acting on the setpoint waveform. The output of the closed loop system is matched to the sum of the outputs of the individual narrow bandpass filters by adjusting the frequency response weight at each narrow bandpass filter in a least squares sense. A block diagram of the schematic is shown in Figure 1.16. The magnitude frequency response of H_1 and H_3 are shown in Figure 1.17.

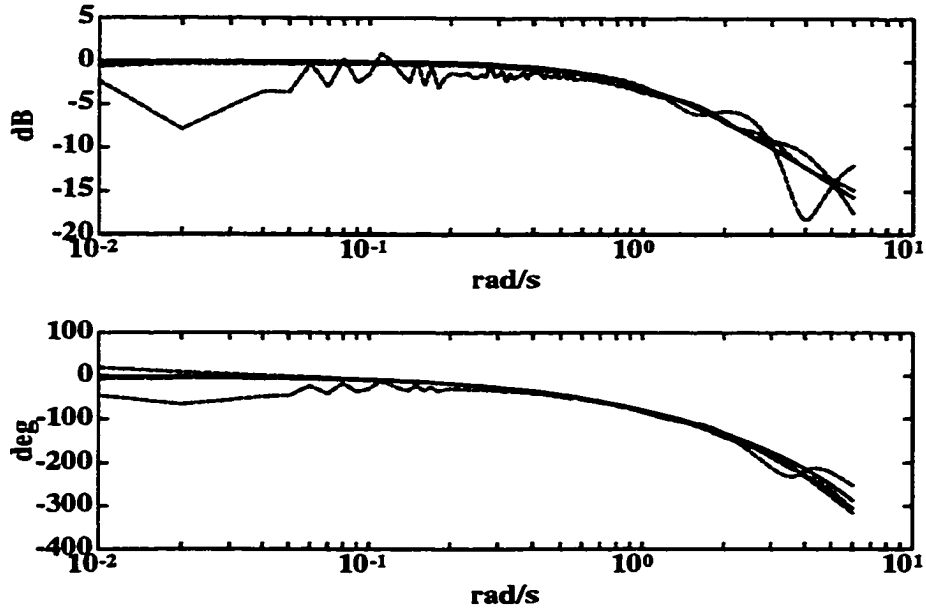


Figure 1.13: Frequency response of process computed from Laguerre identification of closed loop with $p = 2$ and $N = 8, 10, 12$.

Considerable simplifications are possible from the general structure of Figure 1.16. Firstly, since the impulse response of systems are real, the complex frequency response has complex conjugate symmetry. Exploiting this reduces the number of filters in Figure 1.16 to about half the number shown. However, the filters involve complex numbers and therefore the computational burden is still of order N for filtering the setpoint signal. The class of the setpoint signal being applied is usually known such as square/trapezoidal etc. waveform. Consequently the excitation's period and harmonic content is known. For example for a symmetric square waveform, odd harmonics are present and the harmonic content goes down with the harmonic number. This leads to practical upper limit up to which signal content is available after filtering through the FSF channels and it is reasonable to use only the odd harmonics up to this order in computing the FSF channels. In the simulations, with a 30s period square waveform and a sampling period of 50 ms (20Hz), the N of the FSF is 600 (period $\times$ sampling frequency). Using odd harmonics alone in the FSF and the complex conjugate symmetry, the number of channels within the FSF reduces to 150 from 600. Furthermore, since we are only interested in the continuous time frequency response of a process and this is reasonably well approximated by the discrete-time frequency response up to about 6rad/s which is approximately 0.05 of the sampling rate of 20Hz (125.7 rad/sec), computing up to the harmonic number $k = 29$ suffices to reach this frequency. While the signal content in the output and input have higher harmonics present, their amplitudes would be smaller than 1/29 of the fundamental and therefore possibly hidden in the noise in the process output. Neglecting these higher harmonics should not introduce any significant errors in the RLS estimation which uses

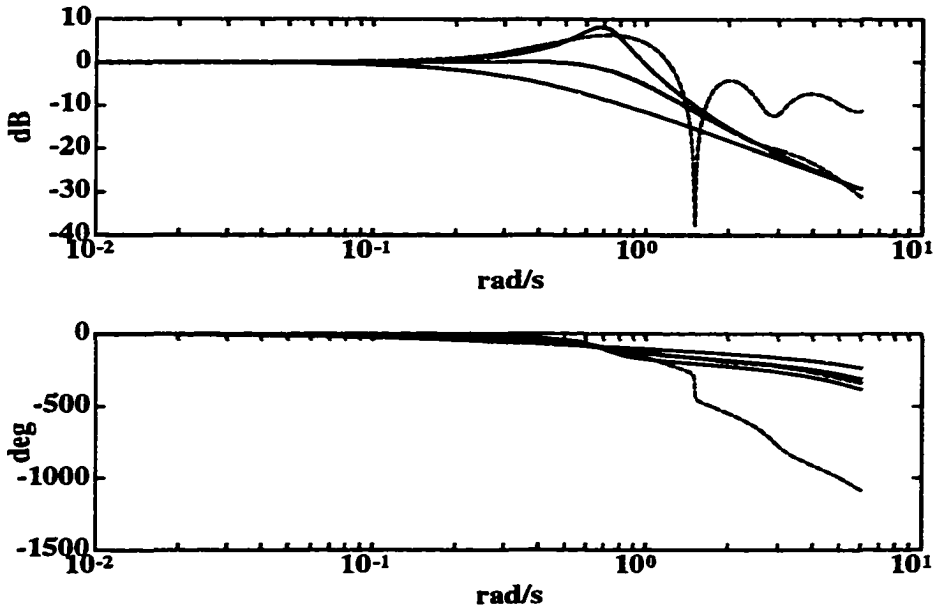


Figure 1.14: Frequency response of process computed from Laguerre identification of closed loop with $p = 2$ and $N = 10$ with the process parameters at nominal and the two extreme sets of process parameters.

the summed outputs from the FSF channels to generate the predicted output of the process (Figure 1.16).

Equation (1.16), like the Laguerre model (Equation (1.12)), allows for interpolation between frequency samples using the filter functions (H_k, Equation (1.16)). However, with the modifications introduced in the FSF, it is not possible to interpolate as the frequency response at even harmonics are not available. Hence with the modifications introduced, the frequency response estimates provided by this method are discrete and finite in number. The estimates obtained on the nominal process and PI controller of this chapter (Equations (1.1, 1.2, 1.3)) after 3000 samples (without disturbance), by the FSF method subject to the modifications discussed above is shown in Figure 1.18. The performance of the modified algorithm in the presence of disturbance is shown in Figure 1.19. The response shows the estimates after one period of excitation (600 samples at 50 ms sampling period) after the disturbance enters the process. From Figure 1.17, it is seen that the magnitude response of each of the interpolating functions $H_k(z)$ introduces exact zeroes at the other centre frequencies and have significant non-zero magnitudes at locations near these frequencies. An alternative way of looking at each $H_k(z)$ is that it introduces a pole-zero cancellation at its center frequency and blocks all other frequencies by introducing exact zeros at the other harmonic frequencies. Such filters are usually sensitive to the excitation frequency mismatch. In identification experiments in the process control industries setpoint bumps (excitations) are usually applied from the operator console and the time frame involved is often of the order of minutes to hours. Exact bump

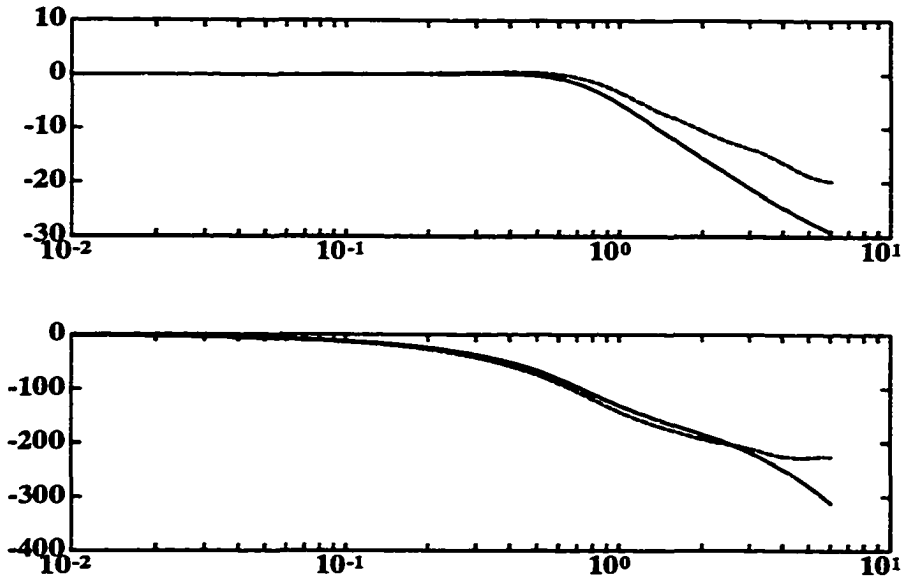


Figure 1.15: Frequency response of closed loop identification of Laguerre model of nominal process in the presence of disturbance.

frequency matched to the FSF design for the identication is therefore almost always impossible. To study the effect of mismatch of excitation frequency on the FSF identification scheme, a simulation was conducted with the nominal process and controller with no disturbance. However, in this simulation the excitation was applied at a 32s period while the FSF scheme was setup assuming that the excitation period was 30s. This constitutes an approximately 7% mismatch in frequency. The performance of the FSF scheme in this simulation after 5 cycles of the excitation is shown in Figure 1.20. The sensitivity of the scheme to excitation frequency is seen in this Figure. The performance of the FSF scheme in identifying closed loop frequency response at the boundaries of process parameter variations and with the nominal process parameters along with the nominal PI controller (Equation (1.3)) is shown in Figure 1.21.

When the FSF scheme performs well in estimating the closed loop, the finite frequency response data of the process computed from closed loop data matches the process frequency response accurately. Therefore no open loop process frequency response plots are provided for this scheme. From these simulations the following conclusions can be drawn:

1. The FSF scheme assumes a finite impulse response model of the process and estimates it. In as much as this order is chosen large enough that the residual tail of the closed loop infinite impulse response is negligible, the method can be applied for process identification.

2. For periodic excitations the method can be modified as per the above to reduce its

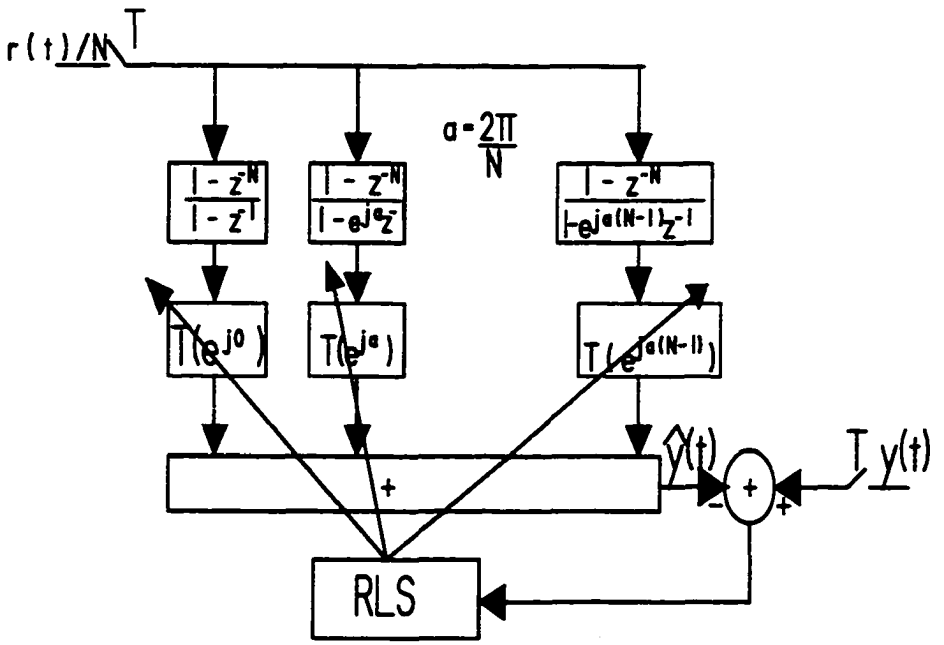


Figure 1.16: Closed loop identification scheme using frequency sampling filter.

computational complexity.

3. Disturbance effects must be averaged out before using the filter's estimates.

4. Frequency of excitation must be well matched to the frequency of excitation assumed in setting up the filter scheme. Otherwise the estimation significantly degrades.

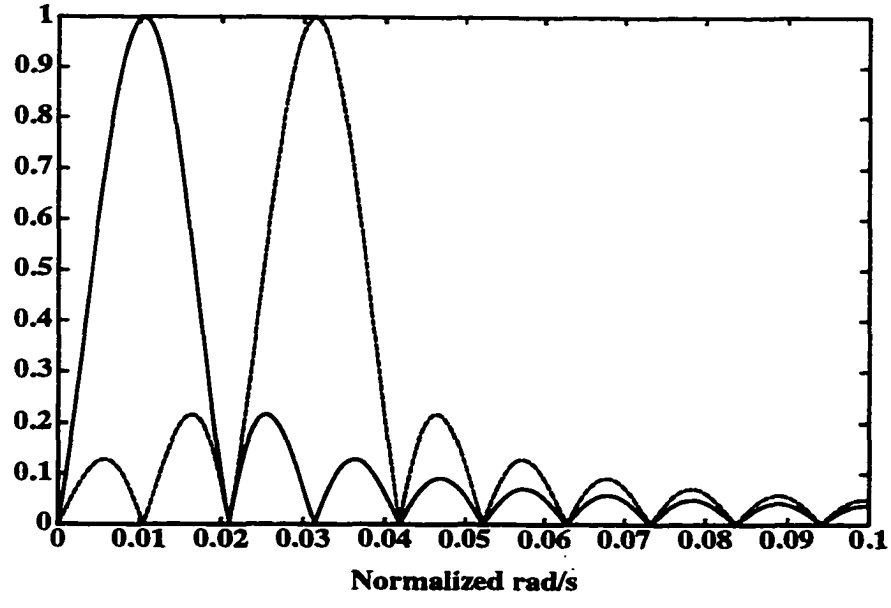


Figure 1.17: Magnitude frequency response of two channels (H_1, H_3) of the frequency sampling filter.

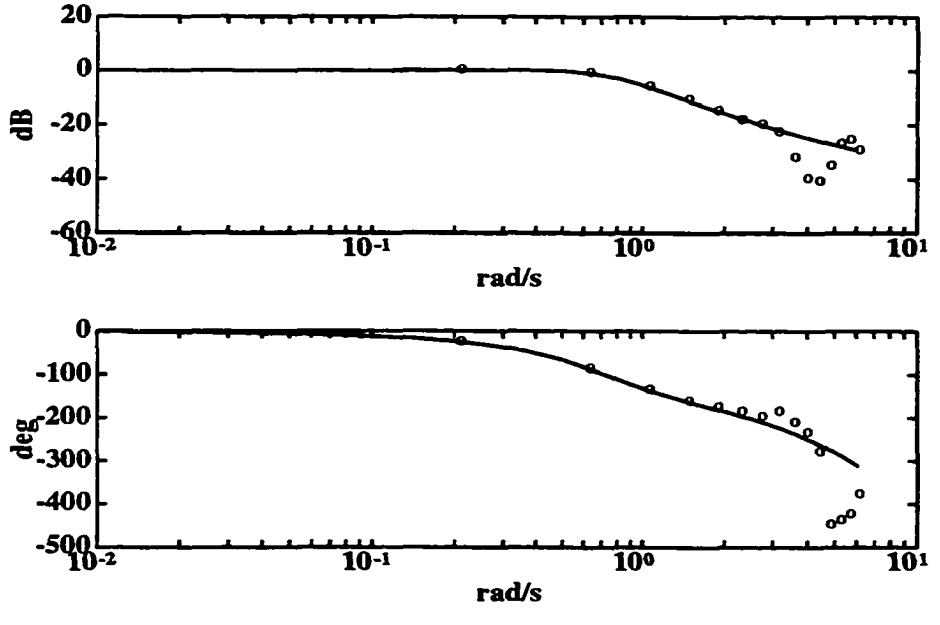


Figure 1.18: Frequency response of closed loop identification using frequency sampling filter to estimate odd harmonic response up to harmonic number 29.

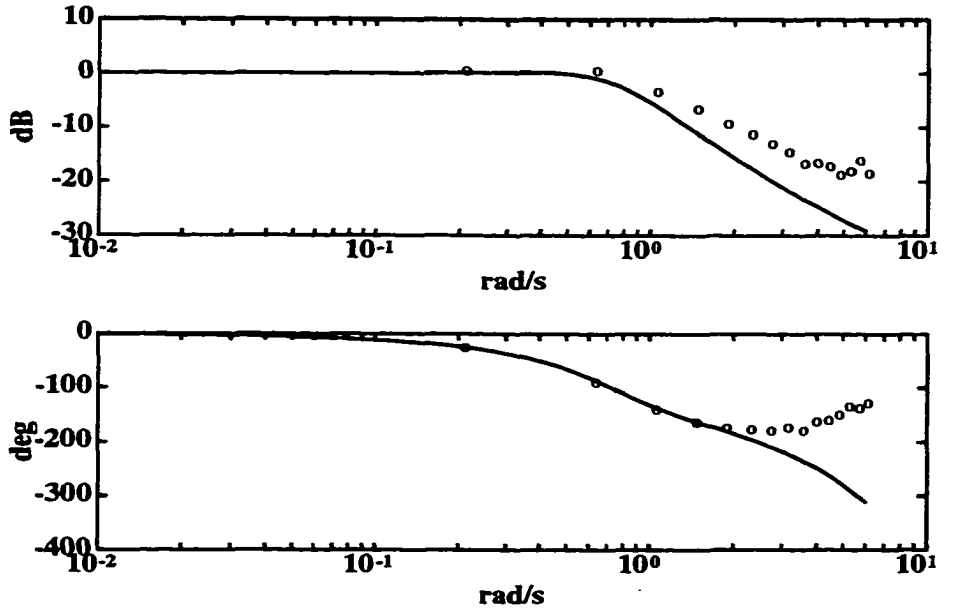


Figure 1.19: Frequency response of closed loop identification using frequency sampling filter to estimate odd harmonic response up to harmonic number 29 in the presence of disturbance.

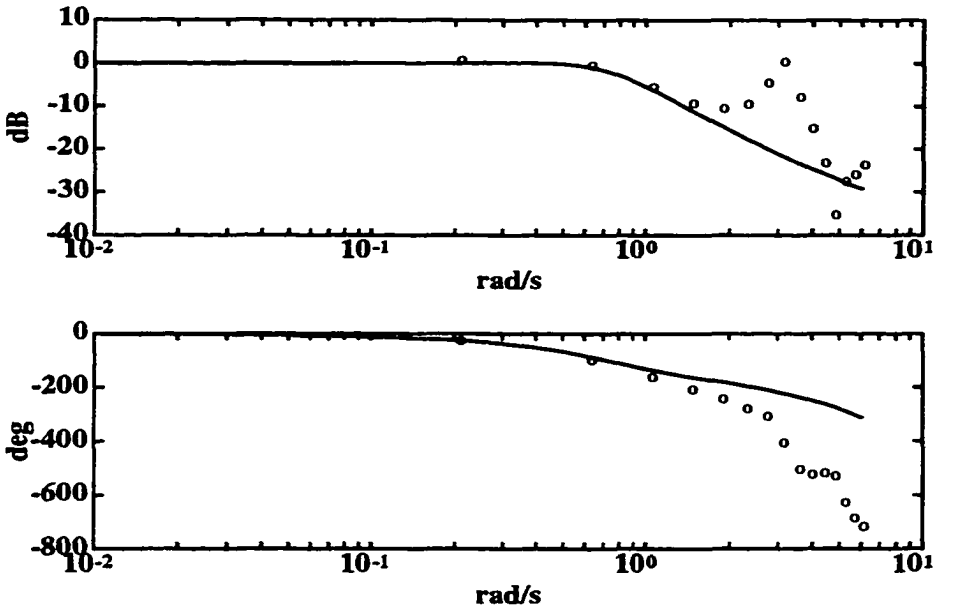


Figure 1.20: Frequency response of nominal process as identified by FSF scheme subject to mismatch in excitation frequency from that assumed in FSF design

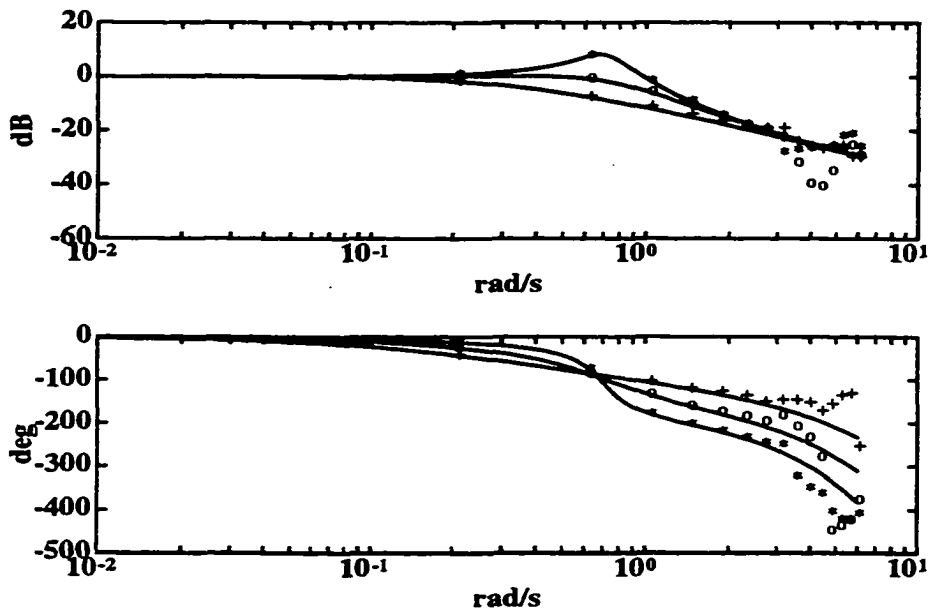


Figure 1.21: Closed loop frequency response of nominal process and process operating at its extreme parameter values as identified by FSF scheme.

1.5 Closed Loop Identification with Bandpass Filters

The bandpass filter based system identification scheme is studied in [15]. The schematic (called one channel in [15]) of estimating the gain and phase shift on one harmonic frequency betwen the excitation and the corresponding frequency in the output is shown in Figure 1.22. The sampled setpoint and the sampled output are fed to a series of parallel channels each peforming frequency response estimation at one frequency point. In every channel, the sampled setpoint and output are fed to typically fourth order Butterworth bandpass (digital) filters. The output of the input bandpass filter is phase-shifted by 90° through a quadrature filter (which is first order in [15]). The recursive least squares algorithm adjusts the coefficients C_1, C_2 until the output of the bandpass filter on the output matches the sum of the weighted output of the input bandpass and quadrature filters. The method also has gaurd bandpass filters which have their center frequencies offset from the main bandpass filters on either side. These filters process a frequency region hopefully devoid of excitation but which have noise components. They thus provide an estimate of noise entering around the channel's excitation frequency and can be used to develop a measure of signal to noise ratio (SNR) [15]. The method provides the estimate of SNR recursively as the estimation proceeds as shown in the schematic. It has generally been observed [15] that a SNR of at least 5 usually indicates convergence in the estimate. Monitoring the SNR on the input and output in each channel guides one to terminate the identification experiment when an appropriate number (usually upto 10 to 15) of the estimates have converged.

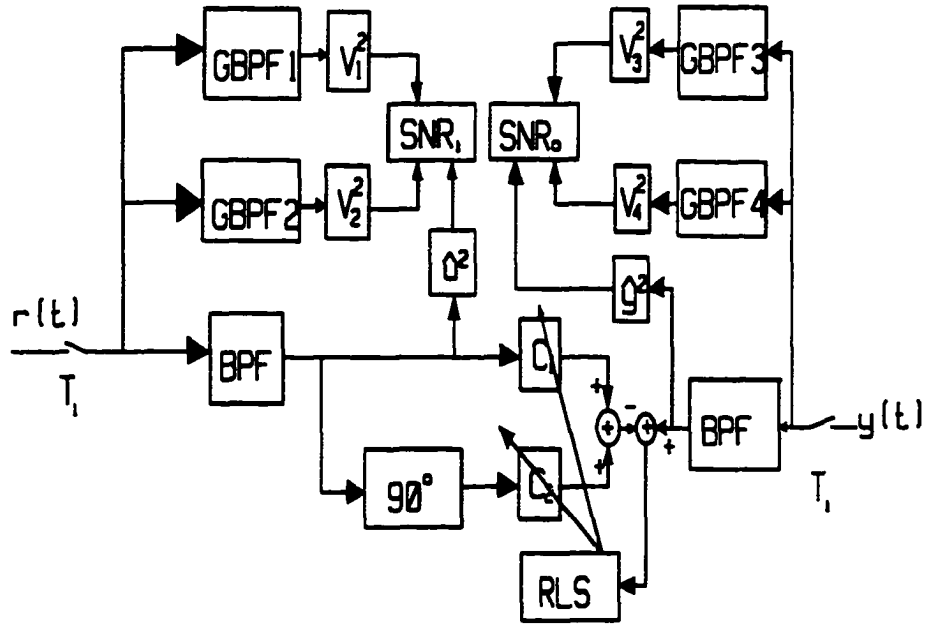


Figure 1.22: Identification Scheme using Bandpass Filters

The plot of the magnitude of fourth order digital butterworth filters centred at the same first and third harmonic frequency as the FSF interpolation functions of Figure 1.17

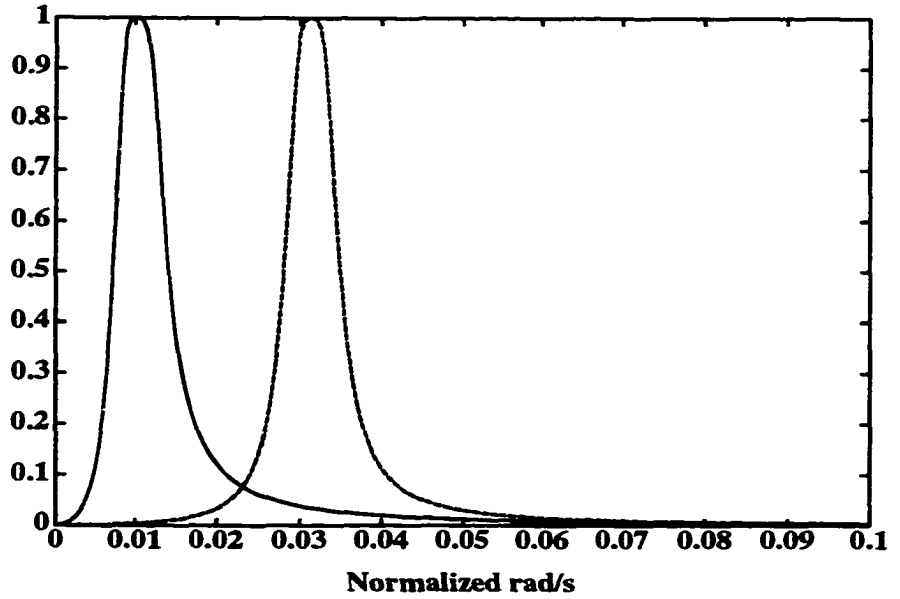


Figure 1.23: Magnitude frequency response of bandpass filters centred at the same frequencies as in Figure 1.17

are shown in Figure 1.23. From this Figure it is seen that the bandpass filters have better attenuation characteristics when one goes away from their center frequency than the FSF interpolation functions.

The key parameters of the parallel bandpass identification scheme (and the values used in simulations reported herein) are: The order of the bandpass filters (4), the center frequency of each filter (odd harmonics of the excitation frequency of 1/30Hz), the -3dB bandwidth of each filter (0.5 of the excitation frequency), the offset of the gaurd filters from the main bandpass filters (0.5 on either side of the center frequency). The RLS parameters for each channel are: initial diagonal covariance (100), forgetting factor for each channel is computed from desired one cycle forgetting factor [15] (0.9) in a period of the harmonic of the excitation passed by the particular channel's filter and initial parameter values (0).

The performance of the bandpass filter identification scheme on the nominal process and controller (Equation (1.3)) is shown in the closed loop frequency response in Figure 1.24. Only 10 points are estimated as the SNR of points beyond the tenth are lower than 3. Consequently, these are deemed unreliable and rejected. The performance of the algorithm when disturbance strikes the process is shown in Figure 1.25. Disturbance affects the performance of the scheme but is of the same order of magnitude as the standard recursive least squares with over parameterized closed loop transfer function (Equation (1.9)). The performance of the algorithm under mismatch in excitation frequency (excitation period = 32s) and the assumed fundamental frequency of the excitation in designing the bandpass filter (assumed period = 30s) is shown in Figure 1.26. The superior performance of the

algorithm in this case to the FSF filter approach is evident. The performance of the

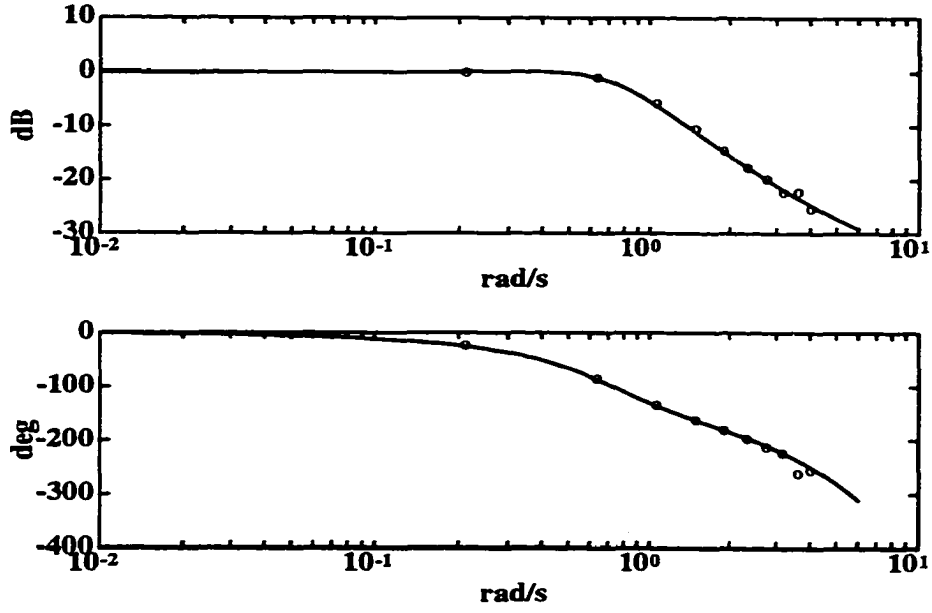


Figure 1.24: Frequency response of closed loop identification using bandpass filters to estimate odd harmonic response upto harmonic number 19.

algorithm for process parameter variations (without disturbance) is shown in Figure 1.27. The performance is akin to the FSF scheme without disturbance.

From these simulations the following conclusions can be drawn:

1. The performance of the bandpass filter identification scheme is comparable with the over parameterized closed loop RLS scheme.

2. The SNR estimates albeit empirically developed in [15] help in establishing reliability of the estimates. Particularly, the upper limit of frequencies upto which the process has been reliably identified.

3. The sampling rate required by this method to adequately control quadrature filter sensitivity errors in this method is shown in [15] to be 20 times the upper frequency upto which identification is sought (assuming noise is smaller than signal strength upto this upper limit harmonic frequency).

4. The main disadvantage of this method is the number of parameters that one needs to set prior to the identification experiment. While real-time code is available in a portable PC with mouse driven menus and graphics at Lakehead University to change the parameters on the fly, bandpass filter order cannot be changed and filter redesign is still not available in the code.

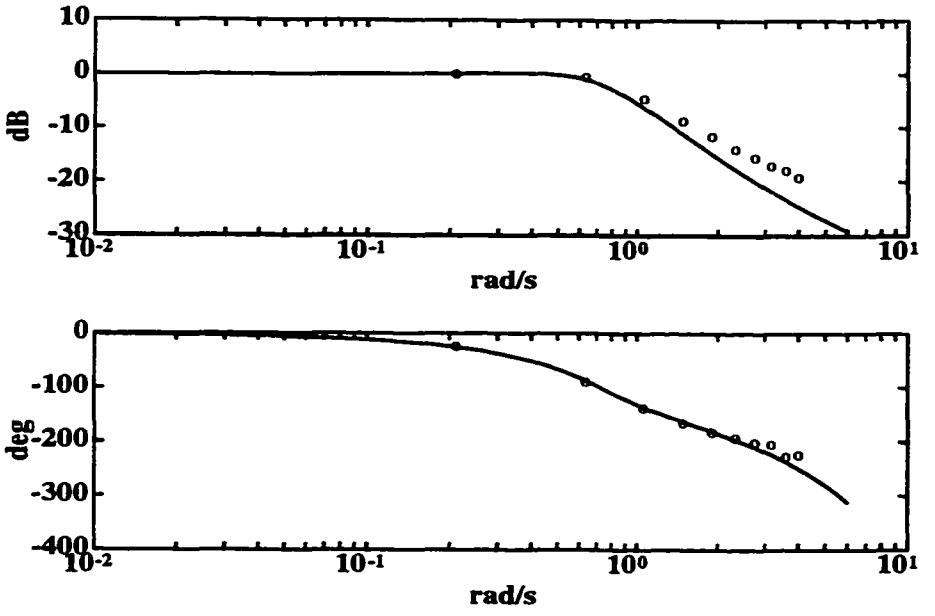


Figure 1.25: Frequency response of closed loop identification using bandpass filters to estimate odd harmonic response upto harmonic number 19 in the presence of disturbance.

1.6 Conclusion

From the simulation study of the various closed loop identification schemes considered in this chapter, it is seen that the most satisfactory schemes of those studied are the over parameterized standard RLS scheme and the bandpass filter scheme.

However, the standard RLS on opening the loop provides reliable estimation of the process frequency response only in the range of frequencies over which the excitation is essentially available. While the lower end of this range is clear for standard RLS, the upper end is usually not explicit. While a similar comment is true for the bandpass scheme, because of the discrete and finite data and the SNR estimation, it provides clearly defined range over which the estimates are reliable. In particular, the upper limit of frequency ranges. In the over parameterized RLS approach, there is still a necessity to estimate the delay "reasonably" in parameterizing the transfer function for the closed loop. The bandpass filter scheme does not use any model but since it measures frequency response needs at least approximately periodic excitation.

In a very definite way, the information on the upper range of frequency upto which the system can be or is identified has to be used in control design in specifying the bandwidth of the desired closed loop for controller tuning. This closed loop bandwidth specification, has to be well within the upper range upto which frequency response data is available. For robust stability, it is also necessary to apply stability margin requirements within this frequency range upto which the estimated data is reliable. The upper range upto which the process is identified imposes a fundamental performance limit. This can be improved only if either the excitation (fundamental) frequency is increased and/or larger amplitude

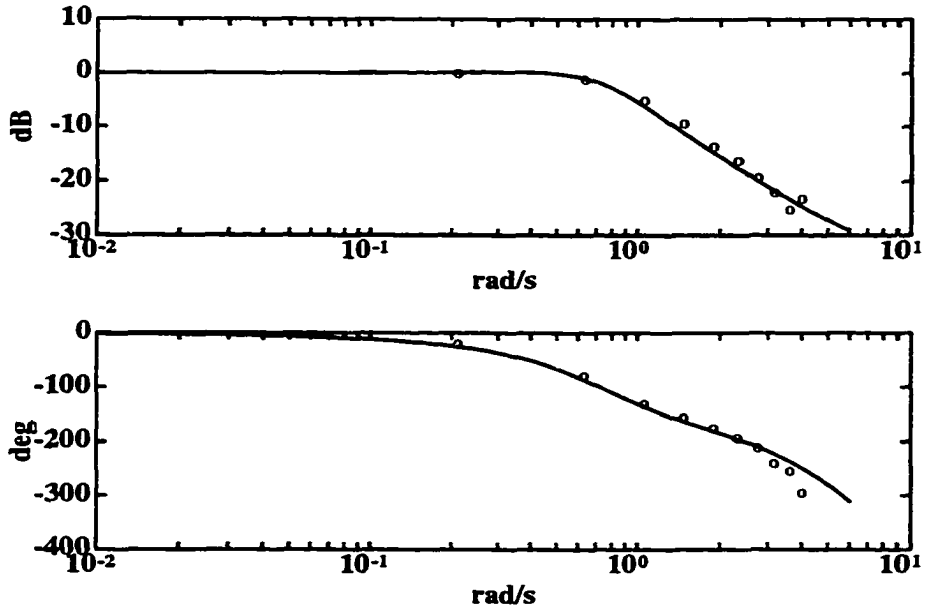


Figure 1.26: Frequency response of nominal process as identified by bandpass filter scheme subject to mismatch in excitation frequency from that assumed in bandpass filter design

of excitation is applied to improve SNR. Large amplitude or high frequency excitations are generally frowned upon in the process control industries where bumps are applied from the operator's console. For example for flow in a pulp mill at 17000 litres/min a 5% bump (100% = 20000 litres/min) implies a sudden flow diversion of 1000 litres/min which usually induces a significant pressure drop in valves somewhere triggering alarms in the process operator's room. Practically, a 2% bump is the atmost that one can expect in such situations. Higher frequency excitation is usually frowned upon as it forces the operator to constantly monitor when to apply the bump and cuts down on his/her time to monitor other points on their console.

In the light of the above, in the further chapters of this thesis the bandpass filter (BPF) scheme is the preferred identification technique.

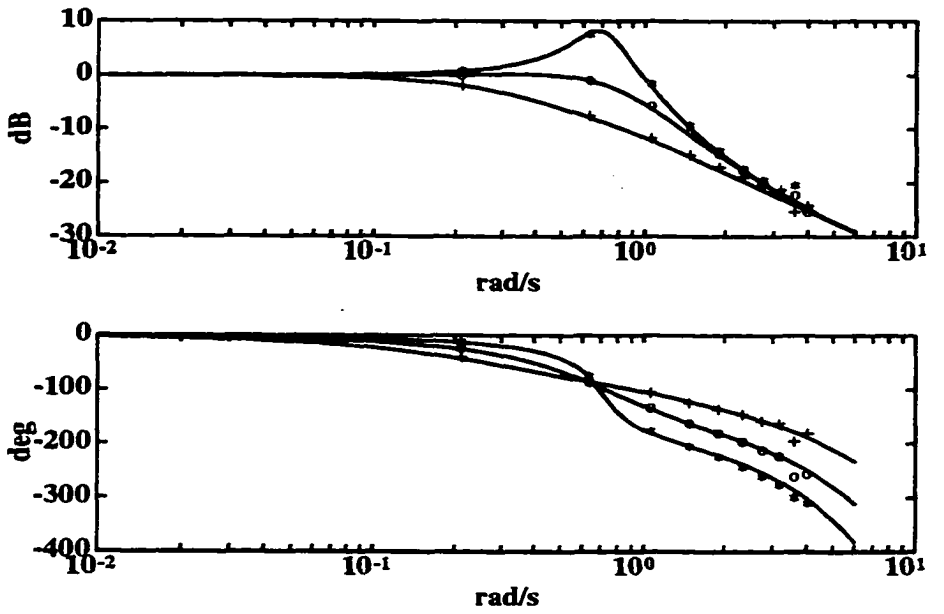


Figure 1.27: Closed loop frequency response of nominal process and process operating at its extreme parameter values as identified by bandpass filter scheme.

Chapter 2

An Overview of Controller Tuning and Design Methods

Controller tuning is the selection of controller parameters to meet a desired robust performance and stability specifications. Controller design methods involve coming up with the structure and the parameters of the controller to satisfy the robust performance and robust stability requirements. In the process control industries most SISO loops are controlled by PID controllers. The tuning of these controllers to effectively reject disturbances under process parameter variations is a key requirement. Explicit MIMO loops are beginning to appear in the process control industries and here too the dominant theme is PID controllers with simple lead/lag or gain type loop decoupling controllers [25].

In this chapter and further chapters of this thesis, it is assumed that the process to be controlled is stable. This is always the case in the process control industry as processes are engineered to be stable. Instability arises usually due to poor controller tuning or process parameter variations from that assumed when the controller was tuned.

There are two categories of controller tuning methods, model based and non-model based. In the model based case, a transfer function of the process is estimated then the controller parameters are obtained. In the non-model based case, the controller parameters are directly obtained based on some minimization criterion and a desired closed loop process response. The latter is more suited to adaptive control while the former is more suited to robust controller designs.

In the sections of this chapter, a quick review of the methods available for controller design and tuning is provided. Then, a particular method [15, 16] to use the finite set of discrete frequency response points obtained from the bandpass filter based identification scheme (Chapter 1, Section 1.5) directly for PID controller tuning for SISO systems is studied. This method is extended for controller tuning to guarantee robust performance over all operating points [21] for which frequency response data is available.

In this chapter the dominant thrust of the literature review is tuning for PID Controllers for SISO systems. In the next chapter when considering MIMO extensions for controller tuning a brief review of some other MIMO methods will be given.

2.1 Kharitonov Theorem Based Methods

In [29] Kharitonov theorem is stated as: "The interval polynomial $p(s,q) = \sum_{i=0}^{n} \left[q_i^-, q_i^+\right] s^i$ is robustly stable if and only if the following four Kharitonov polynomials are stable"

$$k_1 = q_0^- + q_1^- s + q_2^+ s^2 + q_3^+ s^3 + q_4^- s^4 + q_5^- s^5 ..... \quad (2.1)$$

$$k_2 = q_0^+ + q_1^+ s + q_2^- s^2 + q_3^- s^3 + q_4^- s^4 + q_5^+ s^5 ..... \quad (2.2)$$

$$k_3 = q_0^+ + q_1^- s + q_2^- s^2 + q_3^+ s^3 + q_4^+ s^4 + q_5^- s^5 ..... \quad (2.3)$$

$$k_4 = q_0^- + q_1^+ s + q_2^+ s^2 + q_3^- s^3 + q_4^- s^4 + q_5^+ s^5 ..... \quad (2.4)$$

In these Equations, q is the uncertainty vector, and p is the closed loop interval polynomial. Using the sixteen plant theorem [29], the PI controller parameters guaranteeing robust stability for an interval plant family is determined by constructing sixteen Routh tables, that is, one for each Kharitonov plant with a PI compensator given by

$$C(s) = K_p + \frac{K_i}{s} \quad (2.5)$$

The positivity requirement for stability leads to sixteen sets of inequalities, which are given as a function of K_p and K_i . There exist a PI controller to robustly stabilize the interval plant with a set of parameters K_p and K_i , if and only if, there exist a non empty set of gains K_p and K_i satisfying all the inequalities associated with the Routh tables.

The main criticism of the Kharitonov approaches to controller design is the assumption that each of the plant's transfer function coefficients can be anywhere in the interval specified by the uncertainty vector q. In practice, such a random behaviour does not occur, instead the coefficients tend to move in some relations. The relations are not known apriori. Consequently designs using Kharitonov approaches tend to be conservative. Notice also that the thrust of the methods are robust stability rather than robust performance.

2.2 H_2 and H_∞ Methods

The H_2 [32] and H_∞ [31, 30] controllers are robust controllers which minimize the 2-norm or ∞-norm for performance. Either an integral squared error criterion in which disturbance, noise, and uncertainty effects are inherently taken into account is minimized to design the controller, or disturbance, noise and uncertainty effects are taken into account in a linear fractional transformed transfer function matrix [30] and its ∞ norm is minimized to synthesize the controller.

The main problems here are the amount of information needed to capture all the effects into the performance index or the transfer function matrix and the high order controllers that usually result after the design procedure is completed. These controllers are rarely reducible to the standard PID form used in the industry.

While these methods are of tremendous value in the study of achievable performance limits, their spread to the process control industry is not envisaged in the near future.

The reasons being that the following issues (prioritized somewhat arbitrarily from high to low importance) are yet to be considered for these controllers:

1. how to do bumpless transfers between manual and automatic and vice versa on these controllers.
2. the role of these controller during start up.
3. the maintainability of the complex high order controllers by process control engineers in industry (if and when they need tuning) or during a failure and subsequent startup.
4. testing and verifying that the implemented controller works as per the design.

2.3 λ-tuning

The objective in λ tuning [34] is to choose a desired close loop response of a first or a second order type with dead time. The method is based on pole-zero cancellation. The following steps briefly explain the λ tuning procedure for a first order processes with dead time

- find the time constant of the system τ and dead time τ_d
- estimate equivalent time constant $\tau^{'}$,where $\tau^{'} = \tau + \frac{\tau_d}{c}$ and $c = 5 - \frac{\tau_d}{\tau}$
- Tune the PI gains, where $k_c = \frac{\tau^{'}}{\tau}$ and $T_r = \tau^{'}$

The main criticism of this approach is that any method dependent on pole-zero cancellation is sensitive to their variations. Pole-zero variations often occurs in process control systems.

2.4 Ziegler-Nichols Tuning

The controller can be tuned using one of the two methods. In the first method [25], the process output curve shown below in Figure 2.1 can be obtained experimentally by putting the process controller on manual and exciting the system with a bump.

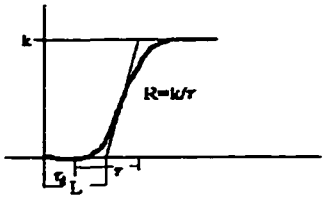


Figure 2.1: Process step response

In this method, the controller parameters are based on a decay ratio of approximately

Type of controller	Optimum gain
P	$k_p = \frac{1}{RL}$
PI	$k_p = \frac{.9}{RL}, T_i = \frac{L}{0.3}$
PID	$k_p = \frac{1.2}{RL}, T_i = 2L, T_d = .5L$

Table 2.1: $C(s) = K_p(1 + \frac{1}{T_i s} + T_d s)$, based on decay ratio of 0.25

0.25 and the parameters are listed in Table 2.1, for P, PI, and PID controllers.

In the second method, the controller parameters are based on driving the system until it is marginally stable. In this case the controller is in the loop but the I and D gains are set to zero. The proportional gain is increased until sustained but non-growing oscillations are achieved and their period is measured as the ultimate period (P_u). The proportional gain at which these oscillations ensue is the ultimate gain (K_u). The P,PI, and PID parameters for controller tuning using this method [19] are shown in Table 2.2 below: Practical implementations using relays of the second method are discussed in [28].

Type of controller	Optimum gain
P	$k_p = 0.5K_u$
PI	$k_p = 0.45K_u, T_i = \frac{1}{1.2P_u}$
PID	$k_p = 0.6K_u, T_i = \frac{P_u}{2}, T_d = \frac{P_u}{8}$

Table 2.2: $C(s) = K_p(1 + \frac{1}{T_i s + T_d s})$, based on stability boundary

These methods are based on practical experience. Any recursive identification of process dynamics in the process control industries is time consuming as processes respond slowly. Consequently since the engineer responsible for tuning has to handle many loops the first method takes far less time than full process dynamics identification. It therefore allows many loops to be tuned in the time available for the engineer. For this reason Shinskey [25] de-emphasizes identification of process dynamics. However these methods are usually not robust to process parameter variations and require retuning when the process is out of tune. During the tuning phase the controller is either on manual or without an I-gain (and therefore cannot guarantee disturbance rejection). Tuning of the controllers using the dynamics of the process always leads to better performance than that obtained from Ziegler-Nichols tuning.

2.5 Dahlin Controller

In the Dahlin controller [33], dead-time compensation is implicitly provided. It is a model-based control method where the desired closed loop response is a first order transfer function with dead-time. The closed loop transfer function is given by

$$G_d(z^{-1}) = z^{-d}\frac{1-p}{1-pz^{-1}} \tag{2.6}$$

The transfer function of the process is assumed to be

$$G(z^{-1}) = z^{-d}\frac{k}{1-\tau z^{-1}} \tag{2.7}$$

Now, by equating the closed-loop transfer function with the desired, the Dahlin controller is obtained as

$$C(z^{-1}) = \frac{(1-p)(1-\tau z^{-1})}{k(1-pz^{-1}-(1-p)z^{-d}} \tag{2.8}$$

The key assumption involved in Equation (2.7) is usually not satisfied in practice and therefore this controller under process parameter variations is not robust.

2.6 PID Tuning Using Finite Frequency Response Data

These PID controller tuning methods in [15, 16, 21, 13] use the system frequency response estimates directly (as provided by the BPF scheme or FSF scheme of process identification discussed in the Chapter 1).

The methods of [15, 16, 21] are based on Fourier series and integral squared error criterion. A desired closed loop second order transfer function with dead time is specified for the performance of the closed loop system. It is assumed to be in the form

$$T_d(s) = \frac{\omega_n^2 e^{-s\tau_d}}{s^2 + 2\xi\omega_n s + \omega_n^2} \tag{2.9}$$

where ξ and ω_n^2 are the design parameters to be specified by the operator to control overshoot and to control speed of response respectively. The dead time in the desired closed loop transfer function is estimated from time domain input-output waveforms obtained during the identification experiment conducted to determine the process frequency response.

The integral squared error criterion in the time domain over one period of excitation, in the steady-state, can be written approximately using Parseval's identity in Fourier Series, as the first (sum) term in the cost function given in Equation (2.11). In Equation (2.11), T_i is the transfer function of the system evaluated at the i^{th} harmonic of the

excitation frequency. It is computed from the corresponding process frequency response estimate G_i and the current candidate PID controller gains. N in Equation (2.11) is the number of reliable frequency response estimates of the process. A_i in Equation (2.11 is the amplitude of the i^{th} harmonic for a conceptual trapezoidal waveform exciting the set point with a period T_{period}, an amplitude of ± 1 units and a slope of K units/sec. A trapezoidal waveform is used instead of a square waveform to reduce the effect of the Gibbs phenomenon in Fourier series of discontinuous waveforms [15, 16]. A_i for the trapezoidal waveform considered, is given by

$$A_i = \frac{4}{\pi}\frac{K}{i^2}\frac{T_{period}}{2\pi}\sin\left(\frac{2\pi i}{KT_{period}}\right) \quad i = 1, 3, 5, ... \tag{2.10}$$

The minimization of the sum of the finite number of terms of the squared error (first sum term) in Equation (2.11) does not guarantee closed loop stability. Hence it is augmented by explicit penalty functions for gain and phase margin violations from desired values. The cost function to be minimized is thus given by

$$F_{Gop} = W\frac{T}{2}\sum_{i=1}^{N}\left(|T_{d_i} - T_i|^2 A_i^2\right) + P_{pm} + P_{gm} \tag{2.11}$$

The penalty functions are chosen as [16, 21]

$$P_{pm}(\phi) = k_1\left[1 + tanh(k_2(\phi_d - \phi))\right] \tag{2.12}$$

$$P_{gm}(m) = k_3\left[1 + tanh(k_4(m_d - m))\right] \tag{2.13}$$

In Equation (2.12), ϕ_d, ϕ are the desired and interpolated phase margins respectively in radians. In Equation (2.13) m_d, m are the desired and interpolated gain margins respectively in dB. k_1, k_2, k_3, and k_4 are positive constants that the designer may choose. W in Equation (2.11) is a positive constant parameter to be chosen by the designer to weight the errors in closed loop performance specifications visavis stability margin violations. Experimental and simulation results of the optimization of this cost criterion by a polytope search method [18] which does not use derivatives are available in [16, 21]. Since the penalty functions are differentiable, optimizations by other techniques which require derivatives are also possible. Experimental and simulation results in [15] use a gradient optimization on the first sum term in Equation (2.11) along with graphical (visual) inspection of gain and phase margin at each iteration. Since the controller parameters enter into T_i (Equation (2.11)) in the numerator and denominator, the optimization is nonlinear. In [13], from the desired closed loop transfer function an equivalent open loop transfer function (GC) is computed by opening the loop. Next, since the controller parameters enter linearly in the open loop function a least squares fit is carried out to obtain the controller parameters.

To obtain a robust controller for a number of operating points, the cost function to be minimized to obtain a single PID controller to provide "good" performance and "good"

stability margins at all operating points can be taken as per Equation (2.14), where l is the number of operating points.

$$F = \sum_{l=1}^{M} W_l F_{G_i} \tag{2.14}$$

2.6.1 Experimental Results

The process considered is an experimental temperature control system whose major blocks are shown in Figure 2.2. The speed of the fan blowing air on the heater is adjustable. The process is noisy and is usually controlled by PI controllers. With a controller

$$C(s) = 4.1 + \frac{2.9}{s} \tag{2.15}$$

and the fan operating at mid-speed, the closed loop frequency response at 6 points was identified by the bandpass filter method. As the process is noisy, only 6 points had a SNR > 5 after 5 periods of excitation. During the identification phase a square wave excitation was applied to the set point with a period of 50s and amplitude of 2.5° C (a 5% bump). The sampling period used in this experiment was 1s. The process delay estimated from the time domain response of the identification experiment was approximately 1s. The bandpass filter method of closed loop frequency response identification (Chapter 1, Section 1.5) was used to identify the closed loop response. From the identified closed loop frequency response, the open loop process frequency response was computed using Equation (1.7). The experiment was repeated with the fan operating at low speed and high speed. The frequency response estimates of the open loop process at the three operating points are shown in Figure 2.3.

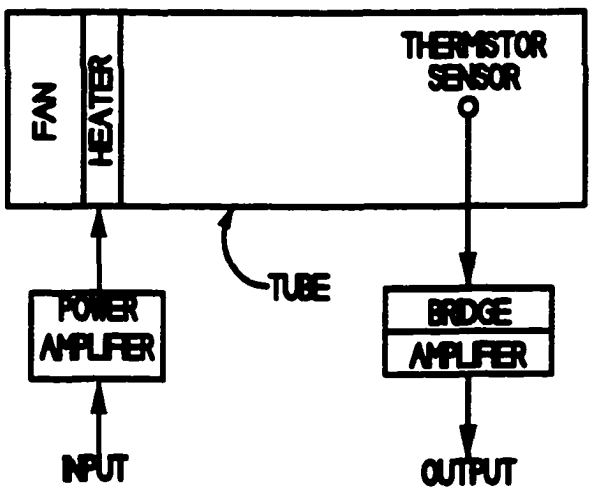


Figure 2.2: Schematic of the temperature control process

Since the controller used in the identification experiment had poor performance at the low fan speed (100% overshoot), the controller was tuned using the technique of [16].

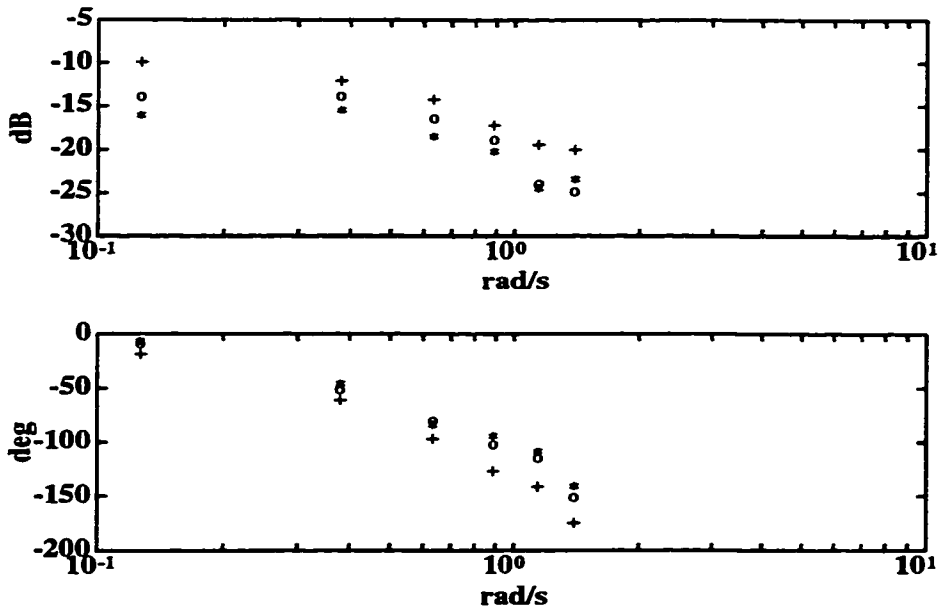


Figure 2.3: Frequency Response Estimates for the open loop temperature process: $+ = G_L$ (fan at low speed), $o = G_M$ (fan at mid-speed), $* = G_H$ (fan at high speed)

The 6 frequency response estimates G_L shown in Figure 2.3 were used. The desired gain and phase margins were 6dB and 35° . The desired closed loop transfer function (Equation (2.9)) was specified with $\xi = 0.707, \omega_n = 0.9$rad/s and delay = 1s. The periodic trapezoidal waveform exciting the set point was assumed to have an amplitude of ± 1 units, a slope of 0.2 units/s and a period of 50s. The optimization using the search method [18] to minimize F_{G_L} (Equation 2.11)) was carried out to obtain a tuned controller given in Equation (2.16). The values of W of Equation in 2.11 was 1 and the $k_i, \quad i = 1, 2, 3, 4$ in the penalty functions of Equations (2.12, 2.13) were set to 50 for this optimization.

$$C_{tuned} = 2.35 + \frac{1.25}{s} \tag{2.16}$$

The time domain performance of the closed loop system with this tuned controller and the process operating with the fan at low speed is shown in Figure 2.4. The tuning has successfully lowered the overshoot.

A robust PI controller to optimize the performance over all three operating points of the process is designed next. The three desired closed loop transfer functions for the three operating points have a delay of 1s and a ω_n of 0.9rad/s. The desired gain and phase margins for each operating point was taken as 6dB and 35° . The damping factor for the closed loop transfer functions (ξ) was taken as 0.9 for G_L, 0.707 for G_M and 0.6 for G_H. The periodic excitation of the set point and the parameters for the cost function (Equation (2.11)) are the same as above. The controller obtained by minimization of the

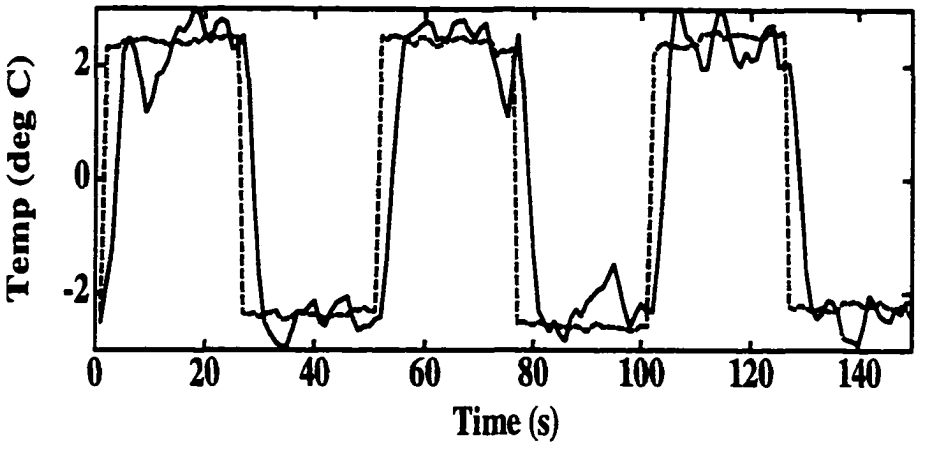


Figure 2.4: Sampled closed loop tuned temperature control system response to a square wave set point excitation with fan at low speed setting

function F in Equation (2.14) is

$$C_{robust} = 2.4 + \frac{1.9}{s} \tag{2.17}$$

The response obtained when this controller was implemented is shown in Figure 2.5.

Attempts to decrease the overshoot at low fan speed makes the response at the high fan speed sluggish. The damping factor specification for each operating point's desired closed loop transfer function has been used to balance these counter effects in this design.

2.7 Conclusion

An overview of controller design methods with emphasis on SISO systems was considered in this chapter. The tuning of controllers using several methods was briefly presented. An extension of the tuning method in [16] to develop a single robust PID controller over many operating points was given and experimental results of the method were provided. The design choices in this method for a SISO system are 4 parameters: the desired closed loop second order transfer function's damping factor to control overshoot, desired natural frequency to control speed of response and desired gain and phase margins for the loop. It is hoped that a designer can intelligently iterate using specifications of these 4 parameters and the tuning technique to arrive at a satisfactory PID controller.

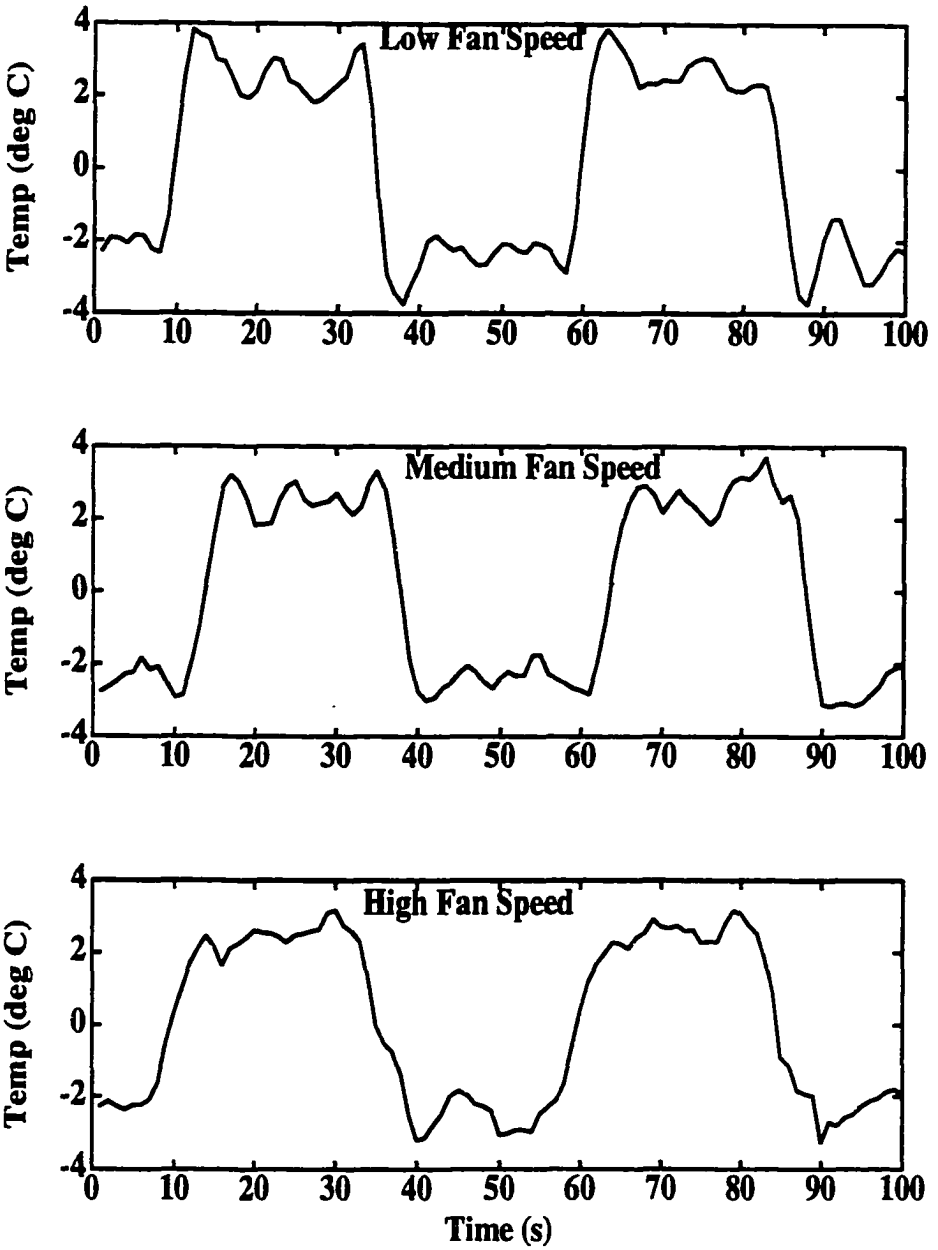


Figure 2.5: Sampled closed loop tuned temperature control system response to a square wave set point excitation with fan at low (top), mid and high(bottom) speed settings

Chapter 3

MIMO PID Controller Design Using Finite Frequency Response Estimates

3.1 Introduction

The SISO PID controller tuning method [16], [21] using finite frequency response estimates for stable systems summarized in the previous chapter uses the phase and gain margin constraints to guarantee stability of the closed loop system. The penalty functions on gain and phase margins (Equations (2.12) and (2.13) respectively) on the process frequency response are differentiable, therefore the controller coefficients can be optimized using any of the optimization techniques such as those requiring differentiability (gradient, conjugate gradient) or non-differentiable such as polytope search (Nelder-Mead [18]). The initial guess for the gains needed by any of the optimization techniques are obtained by either the Ziegler-Nichols rule or from the values used during the closed-loop identification experiment conducted to obtain the open loop process frequency response. The optimization method is used to minimize the weighted least squares error between the desired and actual closed loop responses subject to constraints on gain and phase margins. The desired closed loop response is expressed as the output of a second order + dead time transfer function excited by a periodic wave-form at the set point. The error, at a finite number of harmonics between the desired and the closed loop process frequency responses due to the same excitation are minimized in a least square sense along with penalties for phase and gain margin violations. The number of finite frequencies at which the errors are computed is usually small (5 to 20 harmonics) for typical process control applications. The dead time in the desired closed loop transfer function is obtained from the time domain signals used for estimating process frequency response. The desired second order closed loop transfer function is pragmatically specified by its damping factor and natural frequency. Gain and phase margin requirements for the loop can be specified by rules of thumb.

In MIMO the plant has more than one input and/or more than one output. In general the system will have l inputs and m outputs, however in most feedback systems the number of inputs is equal to the number of outputs. Because there is more than one manipulated variable and more than one controlled variable, interaction among control loops can arise when each manipulated variable affects more than one control variable. Moreover configuring the system i.e. deciding on which input should control what output is no longer simple. The classical relative gain array procedure (RGA) provides the designer with quick means for configuring the system. The steady state gains for each controlled variable relative to each manipulated variable is obtained through differential equations describing the system or experimentally. The elements of the RGA can be arranged in a matrix format and calculated using the steady state gains. Pairing of manipulated variables and controlled variables should be done on those elements of the RGA which are positive and close to one in value. The RGA analysis is based only on the DC gains among different loops in a MIMO system, therefore it does not give the designer a total picture of the system because the plant dynamics are not considered. The modified RGA method described in [25] overcomes the limitations of DC gain analysis and provides the control engineer with means of accommodating the plant dynamics. A brief discussion of the RGA is provided in Section 3.2 of this chapter.

The Inverse Nyquist Array (INA) method is another technique that can be used to analyze a MIMO system for interaction and control design. It is a multivariable frequency response technique that can accommodate plant dynamics for interaction analysis and control design. The objective of the INA technique is to achieve a diagonally dominant system in which the off-diagonal elements of the open-loop plant G are of less important than the diagonal elements. Once a diagonally dominant system is achieved the designer can treat the design of each controller as a single loop problem. a discussion of the INA method is provided in Section 3.2 of this chapter.

The SISO PID tuning technique using finite frequency response described in the previous chapter is extended to MIMO systems in this chapter. The MIMO process is assumed to be stable, linear and time-invariant. The controller is a gain matrix K and a diagonal PID matrix C. The error between setpoints and measurements is fed to the diagonal PID matrix C whose outputs are coupled to the process through the gain matrix K. The frequency response open-loop transfer function matrix of the plant G is identified using any of the identification techniques described in Chapter 1. A penalty function is imposed on the $|I + GKC|$ using the multivariable Nyquist criterion to guarantee stability of the closed loop system. Also minimum singular values of $(I + GKC)$, $(I + (GKC)^{-1})$ and $(I + (KCG)^{-1})$ can be constrained. Since the numerator zeros of the $|I + GKC|$ are amongst the closed-loop poles of the system, they must lie in the left half plane for the system to be stable. The Nyquist criterion is used to determine the stability of the system. The PID controller gains and the decoupling matrix gains are designed simultaneously.

3.2 PID Tuning for MIMO Systems

In MIMO systems, the dimensionality of the tuning problem expands significantly with the number of inputs and outputs to the system. The tuning problem becomes complex, because of the interactions among different loops. Moreover configuring the system is no longer an obvious task for the designer hence, with the increased number of inputs and outputs the question of which manipulated variable should control what output is no longer a simple one. An aid in this analysis is the relative gain method.

3.2.1 Relative Gain Array (RGA)

In [25] a procedure is developed to guide the designer in configuring the system using the RGA method. The relative gain of each controlled variable in a process to each manipulated variable is defined as

$$\lambda_{ij} = \frac{(\frac{\partial y_i}{\partial u_j})_m}{(\frac{\partial y_i}{\partial u_j})_c} \tag{3.1}$$

where i, j are indices to the controlled variable y and the manipulated variable u respectively. Note that λ is the ratio of the change in y_i due to a change in u_j in an open loop effect to that due to a perfect closed loop effect. The numerator and denominator of the above equation are evaluated when all considered loops are on manual (numerator) and when all other loops but the considered loop are on automatic (denominator). The relative gain terms can be arranged in an array as following

$$\Lambda = \begin{array}{c} \\ y_1 \\ y_2 \\ \vdots \\ y_i \\ \vdots \end{array} \begin{array}{c} \begin{array}{ccccc} u_1 & u_2 & \cdots & u_j & \cdots \end{array} \\ \begin{bmatrix} \lambda_{11} & \lambda_{12} & \cdots & \lambda_{1j} & \cdots \\ \lambda_{21} & \lambda_{22} & \cdots & \lambda_{2j} & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ \lambda_{i1} & \lambda_{i2} & \cdots & \lambda_{ij} & \cdots \\ \vdots & \vdots & \vdots & \vdots & \cdots \end{bmatrix} \end{array} \tag{3.2}$$

Relative gain is a dimensionless number therefore it is insensitive to scales, ranges and nonlinearities. One important property of the RGA is that the sum of all terms in a column or a row must add up to unity (a proof of this property is given in [23]).Therefore, the number of terms which must be evaluated is reduced, for example in a 3 by 3 array only four relative gain elements must be evaluated the other elements can be calculated by difference using the above property. The relative gain terms can have a positive or negative value. The relative gain terms can be calculated from a set of steady state gains with all loops open [25]. The open loop steady state gains can be obtained by perturbing each input and measuring the change for all outputs and can be written in matrix form

$$Y = KU \tag{3.3}$$

Now,with $H = K^{-1}$ the relative gain terms are calculated as per

$$\lambda_{ij} = k_{ij}h_{ji} \tag{3.4}$$

where h_{ji} and k_{ij} are the corresponding elements in the K and H matrices. Note that pairing on negative RGA elements should be avoided if possible as such pairing result in unstable system. A detailed discussion of negative RGA elements pairing is given in [23]. The pairing should be done on those elements of λ that are positive and close to 1 in value.

3.2.2 Inverse Nyquist Array (INA) Technique for Controller Design

The INA technique [22] considers the system shown in Figure 3.1 below, in which $G(s)$ is an $m \times l$ plant transfer function matrix, and $K(s)$ is an $l \times k$ matrix representing the input compensator, and $L(s)$ is a $k \times m$ matrix representing the output compensator. A feedback gain matrix F is inserted in the return path, and is assumed to be a diagonal $k \times k$ matrix and is designed so that the closed loop has "good" properties. The transfer

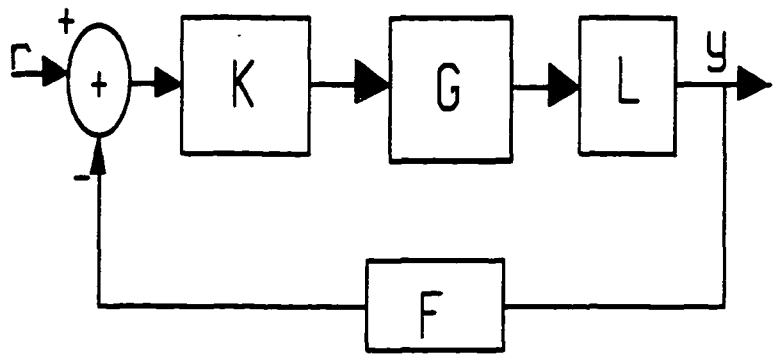


Figure 3.1: INA multivariable closed loop system

function matrix $H(s)$ relating the outputs $y(s)$ to inputs $r(s)$ is given by

$$H = (I + LGKF)^{-1}LGK \tag{3.5}$$

If we write $Q = LGK$ for the transfer function matrix in the input path, then $H(s)$ can be written as

$$H = (I + QF)^{-1}Q \tag{3.6}$$

Now, with $Q^{-1} = Q^{\wedge}$, $H^{-1} = H^{\wedge}$, the closed loop transfer function matrix can be written in a simpler form as

$$H^{\wedge} = Q^{\wedge} + F \tag{3.7}$$

The determinant of $(I + LGKF)$ can be written as

$$|I + LGKF| = \frac{\prod_{i=1}^{n}(s - \beta_i)}{\prod_{j=1}^{p}(s - \alpha_j)} \tag{3.8}$$

where β_i are the closed loop poles of the MIMO system, and α_j are the open loop poles (multivariable) of the system arising from the matrices L, G, K and F. The stability problem of the MIMO system can be defined in the frequency domain using the INA criterion. Writing the determinant of the return difference matrix as

$$|I + QF| = \frac{|H^\wedge|}{|Q^\wedge|} \tag{3.9}$$

The following theorem given in [22] determine the stability of MIMO systems: Let the open loop system have p_o poles in the right half plane: that is, let L, G, K, F have p_o poles (multivariable) there. Then the closed loop system shown in Figure 3.1 is asymptotically stable if and only if

$$N_Q^\wedge - N_H^\wedge = p_o \tag{3.10}$$

Where $N_Q^\wedge$ is the number of clockwise encirclements of $Q^\wedge$ around the origin and $N_H^\wedge$ is the number of clockwise encirclements of $H^\wedge$ around the origin.

If the matrix Q and F are diagonal then the MIMO system becomes a number of SISO loops in which interaction does not exist among the loops, however this situation is limited and too extreme for general cases. The criterion of diagonal dominance overcomes the above limitation and enable the designer to treat a MIMO system as a number of SISO systems. Given a rational $k \times k$ matrix, $Z(s)$, and a contour D in s-plane, if for each $s \in D$

$$|z_{ii}(s)| - \sum_{j=1, j\neq i}^{k} |z_{ij}| > 0 \qquad i = 1, 2 \cdots k \tag{3.11}$$

then $Z(s)$ is row diagonally dominant. Diagonal column dominance is defined similarly for each $s \in D$ by

$$|z_{ii}(s)| - \sum_{j=1, j\neq i}^{k} |z_{ji}| > 0 \qquad i = 1, 2 \cdots k \tag{3.12}$$

If each row or column of $z(s)$ is row or column dominant then $Z(s)$ is diagonally dominant. The Gershgorin bands can be used to check the dominance of a rational function $Z(s)$ as following: First, for each $s \in D$ compute

$$d_i(s) = \sum_{j=1, j\neq i}^{k} |z_{ij}(s)| \tag{3.13}$$

second, for each $z_{ii}(s)$ plot a circle with radius d_i. Third, if each of the Gershgorin bands excludes the origin, for $i = 1, 2, \cdots k$, then Z is row dominant on D. To check for column

dominance one should compute the Gershgorin bands given by

$$d_i'(s) = \sum_{j=1, j\neq i}^{k} |z_{ji}(s)| \tag{3.14}$$

and follow the same procedure described above. The stability of a dominant transfer function matrix for a MIMO system can be determined by checking the diagonal elements only as given in the following theorem [22]:

Let each of the Gershgorin bands based on the diagonal elements $q_{ii}^{\wedge}$ of a diagonally dominant $Q^{\wedge}$ exclude the origin and the point $(-f_i, 0)$. Let these bands encircle the origin $N_{qi}^{\wedge}$ times and encircle the point$(-f_i, 0)$, $N_{hi}^{\wedge}$ times. Then the closed loop system is asymptotically stable if and only if

$$\sum_{i=1}^{k} N_{qi}^{\wedge} - \sum_{i=1}^{k} N_{hi}^{\wedge} = p_o \tag{3.15}$$

Now, the question remaining is how diagonal dominance can be achieved? Assuming $K(s)$ is a stable precompensator applied before the process inputs, and the determinant of $K(s)$ has zeros in the left half plane, one can use $K^{\wedge}$ and $G^{\wedge}$ to make $Q^{\wedge} = K^{\wedge}G^{\wedge}$ diagonally dominant. One way of achieving diagonal dominance is to consider the j^{th} row of $Q^{\wedge} = K^{\wedge}G^{\wedge}$ at some frequency ω.

$$q_{jk}^{\wedge}(j\omega) = \sum_{i=1}^{m} k_{ji}^{\wedge} g_{ik}^{\wedge}(j\omega) = \sum k_{ji}^{\wedge}(\alpha_{ik} + j\beta_{ik}) \tag{3.16}$$

the $k_{ji}^{\wedge}$ for $i = 1, 2, \cdots m$ can be chosen such that

$$\sum_{k=1, j\neq k}^{m} |q_{jk}^{\wedge}(j\omega)|^2 \tag{3.17}$$

is minimized subject to the constraint that

$$\sum_{i=1}^{m} k_{ji}^{\wedge 2} = 1 \tag{3.18}$$

is satisfied.

3.2.3 Singular Value Analysis

The singular value analysis attempts to develop Bode plots for MIMO systems through the use of matrix norms. They describe the magnitude of the loop gain in sensitive directions at each frequency. Given a complex matrix A, the maximum and minimum singular value of A are defined by

$$\overline{\sigma}[A] = \sqrt{\lambda_{max}(A^*A)} \tag{3.19}$$

$$\underline{\sigma}[A] = \sqrt{\lambda_{min}(A^*A)} \tag{3.20}$$

where $\lambda[.]$ are the eigenvalues, and $*$ refer to conjugate transpose. The matrix A^*A is Hermitian with eigenvalues that are real and positive. The singular values of A are the square roots of the eigenvalues of A^*A. The matrix A can be written as

$$A = U \sum V^* \tag{3.21}$$

U and V are an $n \times n$ unitary matrix, and $\sum$ is an $n \times n$ diagonal matrix whose elements are the singular values of A.

The feedback requirement for the design of MIMO unity feedback control systems with a plant $G(s)$ and controller $K(s)$ using singular values [20] is to make

$$\underline{\sigma}[(I + G(j\omega)K(j\omega))^{-1}] \tag{3.22}$$

as large as possible to guarantee robustness to multivariable uncertanties on the output and/or

$$\underline{\sigma}[(I + (K(j\omega)G(j\omega))^{-1}] \tag{3.23}$$

as large as possible to guarantee robustness to multivariable uncertainties on the input over the required range of frequencies. The $\underline{\sigma}(I + GK)$ function is a complementary function to Equation (3.22) [26], and gives the minimum return difference magnitude of the closed loop system. In [26], this singular value function is used to derive concepts of MIMO gain margin.

While in [20] it is shown that performance robustness can also be imposed by singular value analysis, it is not used in this thesis as we specify the desired closed loop transfer function matrix with a specific controller structure for performance robustness as will be discussed in Section 3.3 below.

The singular value analysis enables the designer to analyze a MIMO system in its true multivariable sense without (SISO) loop by loop analysis. The singular values are "equivalent" measures of SISO concepts for the MIMO case. For singular value analysis of robustness to have meaning, **stability has to be ensured for the nominal feedback system [20, 26].**

3.2.4 INA Disadvantage

In the INA method the MIMO problem is reduced to a sequence of SISO problems that are manipulated independently using classical techniques. In [20] it is shown that while the INA set of SISO systems might have a good robustness properties, the resulting closed loop MIMO system may have poor stability robustness properties when analyzed with singular values. The PID tuning using finite frequency response technique described in Section 3.3 overcomes the disadvantages of the INA by taking the singular values of the system in to account in the design and the decoupling gain matrix K and the diagonal PID controller gains are simultanously designed.

3.3 PID Tuning for MIMO Systems using Finite Frequency Response

The block diagram for the described closed loop system is shown in Figure 3.2 below. The number of outputs, number of inputs, and number of setpoints are assumed to be equal. The PID controller C is assumed to be a diagonal matrix. The decoupling gain matrix

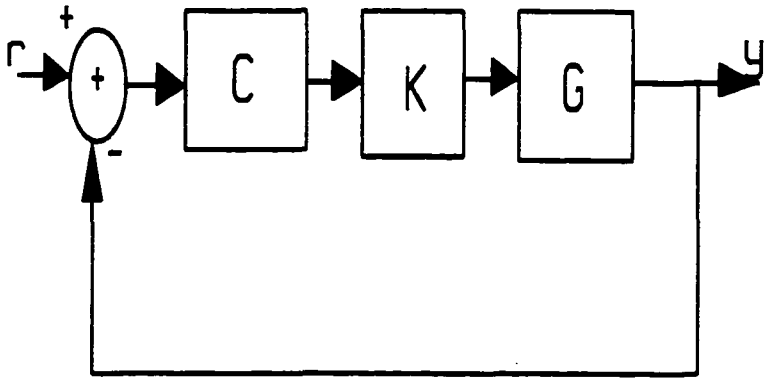


Figure 3.2: MIMO PID + gain control system configuration

K is imposed between the PID controller output and the process input. The controller configuration and connections to INA is obvious. This configuration ensures that the setpoints tracks the measured process outputs.

The plant is assumed to be stable, linear and time-invariant. The finite number of frequency response points of the process required for tuning the MIMO controller are assumed to be spaced at harmonics of a periodic waveform exciting the setpoint of the closed loop system. These points maybe directly obtained by frequency domain identification methods described in Chapter 1 such as the bandpass filter approach, the FSF method etc.. The system is closed with a unity feedback. For tuning purposes, the system is assumed to be excited by a periodic trapezoidal setpoint (vector) waveform with a period of T_{period}, a swing of ± 1, and a slope of K units per sec. The Fourier series coefficients for an odd symmetric trapezoidal waveform at harmonic number i of the fundamental frequency $f = \frac{1}{T_{period}}$ for an input k are given by

$$A_i^k = \frac{4K}{\pi i^2} \frac{T_{period}}{2\pi} \sin\left(\frac{2\pi i}{KT_{period}}\right) \qquad i = 1, 3, 5, \cdots \tag{3.24}$$

The error between a desired closed loop output response to the input waveform and the obtainable closed loop response for the process is minimized in a least squares sense in the frequency domain to tune the multivariable controller. The desired closed loop response is specified as a diagonal matrix with diagonal elements being second order + dead time transfer function acting on the input setpoint waveform. The matrices K and C are simultaneously designed using the frequency response estimates of the process at a finite

number of frequencies, to nudge the closed loop transfer function towards the desired closed loop with penalties on deviation from stability margins. These penalties are imposed on singular value deviations from desired singular values of appropriate frequency response matrices [20], in addition to imposing stability requirements assessed using the multivariable Nyquist criterion on an equivalent open loop SISO transfer function computed as $|I + GKC| - 1$ subject to modifications discussed in Section 3.3.1. While gain margin and phase margin requirements may be imposed on $|I+GKC|-1$, examples such as in [26], show that they may not have significance and therefore singular value analysis must be always carried out after guaranteeing stability for the closed loop. The function to be minimized in the MIMO case at a particular operating point is thus given by

$$F_{G_{op}} = W\frac{T}{2}\sum_{k}\sum_{j}\left[\sum_{i=1}^{N}\left(|T_{d_i}^{j,k} - T_i^{j,k}|^2\{A_i^k\}^2\right)\right] + P_{sv} + P_{Nyquist} \tag{3.25}$$

where the j, k superscripts are indices addressing the outputs and inputs respectively and i refers to the i^{th} harmonic frequency of the (k^{th}) set point excitation. $p_{Nyquist}$ is the penalty function on stability violations and is discussed in Sections 3.3.1, 3.3.2 and 3.3.3. p_{sv} is the penalty function on singular values and is discussed in Section 3.3.3. The desired closed loop transfer function matrix T_{d_i} is a diagonal matrix chosen by the designer in order to satisfy a desired performance.For a 2-input 2-output system it can be written as.

$$T_d(s) = \begin{bmatrix} \frac{\omega_{n_1}^2 e^{-\tau_{d_1} s}}{s^2+2\zeta_1\omega_{n_1}+\omega_{n_1}^2} & 0 \\ 0 & \frac{\omega_{n_2}^2 e^{-\tau_{d_2} s}}{s^2+2\zeta_2\omega_{n_2}+\omega_{n_2}^2} \end{bmatrix} \tag{3.26}$$

The desired second order transfer function for each diagonal element is selected according to desired specifications such as bandwidth, overshoot, rise time (and disturbance rejection). The off diagonal elements are chosen to be zero to eliminate interaction among different loops in the MIMO system. $T_i^{j,k}$ is the complex number describing the frequency response of the actual closed loop system matrix at the frequency $\frac{i}{T_{period}}$ for the j, k element of T given by

$$T = (I + GKC)^{-1}(GKC) \tag{3.27}$$

3.3.1 Modification of Integrator Poles

The PID controller for each loop introduces an integrator in $|I+GKC|$, *i.e.* each integral action in a loop will introduce a pole at the origin. For a MIMO system this implies that for a stable closed loop system the $|I + GKC|$ will have at least 2 or more poles at the origin and is therefore open loop unstable. The stability of the MIMO system can be determined using the MIMO Nyquist criterion. Let the Nyquist contour D be indented to the left half plane to include all imaginary poles and let the open loop system have p_o poles in the closed right half plane then the following corollary in [27] describe the stability of a MIMO system. The closed-loop system is asymptotically stable if and only

if the map of D by $|I + GKC|$ encircles the origin $-p_o$ times clockwise. where D is large enough to enclose all finite poles and zeros of $|I + GKC|$ lying in the closed right half-plane. For closed loop stability, therefore in our case, the numerator of $|I + GKC|$ must be stable as the process G is assumed to be stable. When using optimization techniques it is difficult to compute the number of encirclements using finite frequency response to determine stability. Moreover on line computing is almost impossible due to the infinite arc of the map of the contour D by $|I + GKC|$ poles at the origin due to the integrators. To overcome this problem the $|I + GKC|$ is multiplied by a weighting function to cancel the open loop poles at the origin as per Equation (3.28).

$$\text{modified}(|I + GKC|) = |I + GKC|\frac{s^p}{(\tau s + 1)^r} \tag{3.28}$$

In Equation (3.28) p is the power that is needed to cancel the poles at the origin of $|I + GKC|$ and is equal to the number of inputs or outputs. r is the power in the denominator of the weighting function with $r \geq p$. The parameter τ in Equation (3.28) can be chosen beyond the bandwidth of interest. The denominator $(\tau s + 1)^r$ is introduced to ensure that the Nyquist plot is proper.

The modified determinant can be used by subtracting 1, to generate the equivalent transfer function whose traditional open loop finite frequency response should not encircle the -1 point. Note that this does not change the closed loop poles which are amongst the roots of the numerator of $|I + GKC|$. Now counting the number of encirclements to determine the stability of the MIMO system is simplified. If the finite frequency response of the modified $|I+GKC|-1$ encircle the -1 point then the system is not stable, and if the finite frequency response of the modified $|I + GKC| - 1$ do not encircle the -1 point then the system is stable. It is assumed that the finite frequency response data of the process is available to ascertain this. Therefore, in our approach for a stable system, the stability of $|I+GKC|$ can be easily obtained by verifying that the finite frequency response points within the above modification does not encircle the -1 point.

3.3.2 Nyquist Plot Penalty Function

To guarantee stability of the closed loop MIMO system, the Nyquist plot of the modified $|I + GKC| - 1$, must not encircle the -1 point, therefore a penalty function is applied to the finite frequency response data of the modified $|I + GKC| - 1$ and any attempt to encircle the -1 point will be penalized. A wedge shaped area is defined on the Nyquist plot and the finite frequency response points of the modified $|I + GKC| - 1$ are forbidden from entering the wedge area. The wedge area boundaries can be defined by two lines (I_n, I_p) passing through a vertex $(-G_m)$ as given in Figure 3.3. Note that the wedge does not extend to the right of the vertex but is only to its left.

A measure of "gain margin" constraint is directly specified in this plot by the choice of the vertex at $-G_m$. As will be seen later in Chapter 4 this does not have much significance as specifying points such as -0.6 may be impossible to attain with the control

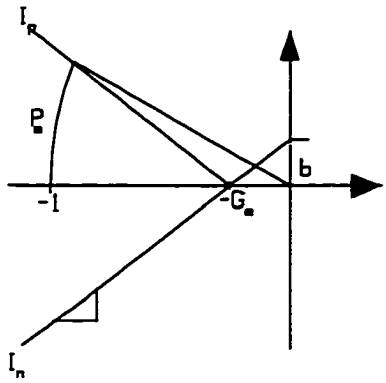


Figure 3.3: Wedge in Nyquist plot

configuration considered herein for practical systems. Also the example in [26] shows such specifications do not guarantee much for MIMO systems unlike SISO systems. A measure of phase margin P_m is indirectly specified on the plot through the slope m of the boundaries of the wedge. The line equations governing the boundaries are specified by the slope m and the intercept b with the imaginary axis as per Equation (3.29).

$$\begin{aligned} m &= \frac{\sin(P_m)}{\cos(P_m) - G_m} \\ b &= mG_m \end{aligned} \tag{3.29}$$

It is assumed that G_m and P_m are chosen so that m and b are meaningful. In order for the MIMO system to be stable the frequency response of the modified $|(I + GKC)| - 1$ should not cross the region bounded by the straight lines forming the wedge. To check if the Nyquist plot of the modified $|(I+GKC)|-1$ passes through the wedge, the imaginary axis points I_p, I_n on the wedge boundaries for each candidate negative real value of the frequency response ($< -G_m$) of the modified $|(I + GKC)| - 1$ are computed. The actual imaginary values for the coresponding frequency points should not lie between the two limit points. $I_p(i)$ and $I_n(i)$ are given in Equation (3.30)

$$\begin{aligned} I_p(i) &= -mRe(\text{modified}((I + GKC) - 1) - b \\ I_n(i) &= mRe(\text{modified}((I + GKC) - 1) + b \end{aligned} \tag{3.30}$$

where Re denotes real part.

Note that if the frequency response points are widely separated then two points can lie on either side of the wedge and the testing will not reveal that the plot actually goes through the wedge. In this case, linear interpolation between the points would catch the situation.

In most applications instead of using the G_m, P_m specifications, it suffices to choose the wedge with a vertex (just) to the right of the -1 point and wedge half angle at 1 or 2° rather than through the formal procedure outlined in this subsection. The equivalent G_m and P_m can be reverse computed, but in the light of their poor value as indicators of stability robustness are of little use.

3.3.3 Penalty functions

The Nelder-Mead polytope search method [18] is used in our work to solve the optimization problem. This search method is a downhill simplex method requiring only function evaluations (no derivative evaluations), therefore the penalty functions need not be differentiable. The Nyquist plot penalty function guarantee the stability of the MIMO system by penalizing the finite frequency response of the modified$|I + GKC| - 1$ if they attempt to encircle the -1 point as given in the Equations below:

$$P_{Nyquist} = 0 \tag{3.31}$$

For each finite frequency response point with negative real part ($< -G_m$) for the modified $|I + GKC| - 1$,

$$IF I_n < Im(\text{modified}(|I + GKC| - 1)) < I_p \quad THEN \quad P_{Nyquist} = P_{Nyquist} + 100000 \tag{3.32}$$

The singular value penalty function guarantee good robustness for the MIMO system. Let the minimum singular values of $I + (GKC)^{-1}$ and $I + (KCG)^{-1}$ at each frequency point be denoted $SV_1(i)$ and $SV_2(i)$ respectively. Now the singular value penalty function P_{sv} can be written as

$$P_{sv} = 0 \tag{3.33}$$

For each finite frequency point

$$IF \quad (SV_1(i) < c_1) \ OR \ (SV_2(i) < c_2) \quad THEN \quad P_{sv} = P_{sv} + 100000 \tag{3.34}$$

c_1 and c_2 are constants to be specified by the designer. Usually c_1 and c_2 are chosen as values between 0 and 1 with values closer to 1 for greater robustness.

3.3.4 Example

To illustrate the MIMO tuning technique described in this chapter, an example is provided here without singular value penalization. Consider the 2-input, 2-output process transfer function of a compressor [22] and desired closed loop transfer function matrices given below

$$G(s) = \begin{bmatrix} \frac{0.1133e^{-0.715}}{1.783s^2+4.48s+1} & \frac{.9222}{2.071s+1} \\ \frac{0.3378e^{-0.299s}}{0.361s^2+1.09s+1} & \frac{-0.321e^{-0.94s}}{0.104s^2+2.463s+1} \end{bmatrix} \tag{3.35}$$

$$T_d(s) = \begin{bmatrix} \frac{e^{-0.8s}}{s^2+2s+1} & 0 \\ 0 & \frac{16e^{-s}}{s^2+8s+16} \end{bmatrix} \tag{3.36}$$

The gain matrix K can be written as

$$K = \begin{bmatrix} 1 & \frac{k_{12}}{k_{22}} \\ \frac{k_{21}}{k_{11}} & 1 \end{bmatrix} \begin{bmatrix} k_{11} & 0 \\ 0 & k_{22} \end{bmatrix} \tag{3.37}$$

By lumping k_{11} and k_{22} with the diagonal PID controller, only the ratios of the off diagonal columns to the main diagonal elements need to be estimated. Therefore, in our example only eight gains need to be designed three for each PID controller and two gain ratios for the decoupling gain matrix K. The system is assumed to be excited at its set points by periodic trapezoidal waveforms with a period of 75s, unit amplitude and slope of 0.1 units/s. 20 odd harmonics of the input excitation and process frequency response are used for minimizing the function in Equation (3.25). For this example the stability margin violation constraints are applied on the modified$|(I + GKC)| - 1$ with a "gain margin" $\left(\frac{1}{G_m}\right)$ requirement of 6 db and "phase margin" P_m of 5° . No constraints on singular values were applied in this example. The Nelder-Mead optimization technique is used to arrive at the results given in Equation (3.38) and Equation (3.39). As is typical of many parameter non-linear optimization problems, the optimal solution depends on the initial conditions and therefore iterations are needed to arrive at a satisfactory design.

$$C(s) = \begin{bmatrix} 0.40 + \frac{0.30}{s} + 0.025s & 0 \\ 0 & 0.25 + \frac{0.2}{s} + 0.00s \end{bmatrix} \tag{3.38}$$

$$K = \begin{bmatrix} 1 & 8.96 \\ 1.14 & -1 \end{bmatrix} \tag{3.39}$$

A plot of the achieved frequency response compared with the desired transfer function is shown in Figure 3.4. The minimum singular values of $(I + GKC)$, $(I + (GKC)^{-1})$ and $(I + (KCG)^{-1})$ for this design are 0.64 , 1.00 and 0.84 respectively indicative of the good stability margins for this design. To assess stability by a different approach, the time delays in Equation (3.37) are approximated by a first order Pade approximation. The transfer function of $|(I + GKC)|$ is calculated and the roots of its numerator and denominator are computed and listed in Table 3.1. From this Table, since the denominator retains its two unstable roots at $s = 0$ and all other roots are stable, it is easily seen that closed loop characteristic polynomial of the MIMO system is stable.

Roots of Numerator of $\|I + GKC\|$	Roots of Denominator of $\|I + GKC\|$
-23.3090	0
-7.4193	0
-2.7979	-23.2695
-0.8541 ±j1.5079	-6.6890
-2.8390	-2.7972 ± j0.0003
-2.1694 ± j0.0720	-2.2650, -2.1276
-0.4142 ± j0.1995	-1.5097 ± j0.7006
-0.2917 ± j0.1259	-0.4132, -0.2476

Table 3.1: Roots of numerator and denominator of $|I + GKC|$

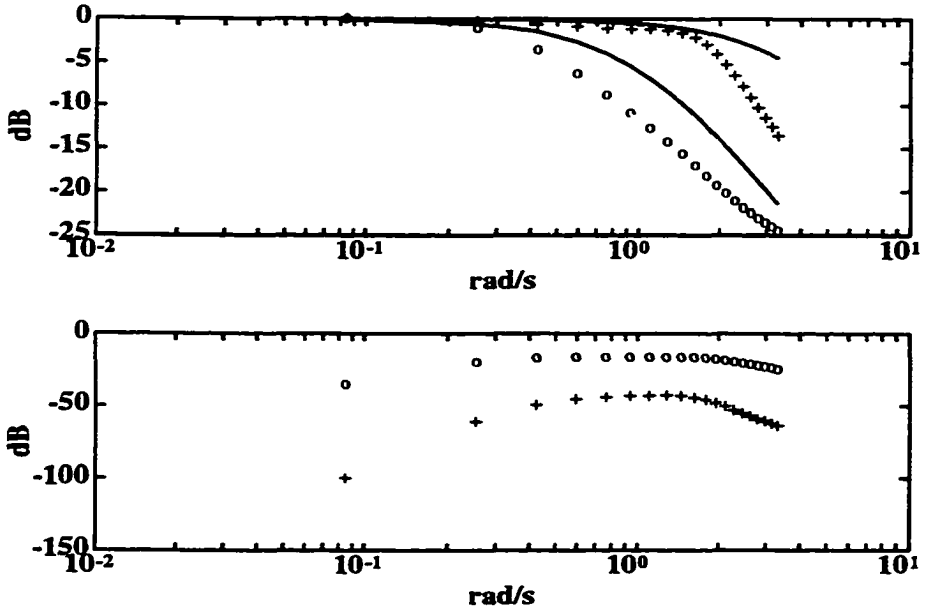


Figure 3.4: Desired and achievable closed loop frequency response for MIMO example: diagonal elements (top) and off-diagonal elements (bottom)

3.4 Practical Issues

In practice sometimes some loops are put on manual while other loops operate on automatic. However doing so could upset a system which is maintained stable by a "true" multivariable controller. Shinskey [25]explains and describes such situations. The penalty function approach described in this section enable the possibility of each main loop being on automatic while the rest of the loops are on manual while maintaining stability (and viceversa). In a 2-input, 2-output system there are two main loops to be considered the first loop on automatic and the second loop on manual and viceversa. If the second loop is on manual and the first loop is on automatic one can calculate the corresponding $|I + GKC| - 1$ and then impose stability margins on its finite frequency response data. Similarly the $|I + GKC| - 1$ can be calculated when the second loop is on automatic and the first loop is put on manual, and stability margins can be imposed on its finite frequency response data. The functions given below describe the $|I + GKC| - 1$ for each possible case.

$$loop(1)_{Automatic} = (g_{11}k_{11} + g_{12}k_{21})c_{11} \tag{3.40}$$

$$loop(2)_{Automatic} = (g_{22}k_{22} + g_{21}k_{12})c_{22} \tag{3.41}$$

Note that in this case $loop(1)$ and $loop(2)$ have only one integrator pole at the origin and are SISO transfer functions and therefore we impose gain and phase margin constraints on the finite frequency response points along the lines of Chapter 2, Equations (2.12, 2.13).

The following function can now be optimized

$$F_m = F_{loop1} + F_{loop2} + F_{G_{op}} \tag{3.42}$$

where F_{loop1} and F_{loop2} are given below. Note that F_m incorporates true MIMO operation plus the case of one loop being on manual and viceversa.

$$F_{loop1} = P_{pm_1} + P_{gm_1} \tag{3.43}$$

$$F_{loop2} = P_{pm_2} + P_{gm_2} \tag{3.44}$$

In Equations (3.43, 3.44), P_{pm_i}, P_{gm_i} are the penalty functions described in Chapter 2, Equations (2.12, 2.13), made specific for phase and gain margin violations respectively using the transfer functions $loop(1)_{automatic}$ and $loop(2)_{automatic}$. $F_{G_{op}}$ is given by Equation (3.25). Note that Equation (3.43) and Equation (3.44) are additional constraints to the MIMO problem.

3.4.1 Example

To illustrate the approach of this section, the same example given in the previous section was used with the addition of the constraints given in Equations (3.43, 3.44). The additional constraints of Equation (3.43) and Equation (3.44) are used with a desired gain margin of 12dB and a desired phase margin of 75° for both possible cases. All other MIMO desired characteristics were the same as in the previous example. The Nelder-Mead optimization technique [18] is used starting with the previous controller parameters (Equations (3.38), (3.39)) as initial guess, to arrive at the results given in Equations (3.45, 3.46).

$$C(s) = \begin{bmatrix} 0.49 + \frac{0.30}{s} + 0.02s & 0 \\ 0 & 0.26 + \frac{0.2}{s} + 0.03s \end{bmatrix} \tag{3.45}$$

$$K = \begin{bmatrix} 1 & 8.91 \\ 1.19 & -1 \end{bmatrix} \tag{3.46}$$

A plot of the achieved frequency response compared with the desired transfer function is shown in Figure 3.5. The minimum singular values of $(I + GKC)$, $(I + (GKC)^{-1})$ and $(I + (KCG)^{-1})$ for this design are 0.71, 1.00 and 0.96 respectively indicative of the good stability margins for this design. In comparison with Figure 3.4, Figure 3.5 reveals that the off-diagonal elements have risen up in the new optimization and that the 2,2 diagonal element of the closed loop ('+' points) now do not match the desired as well as before. This is intuitively correct as stringent individual loop stability margins have been imposed in this example. Notice also that this is borne out by the improvement in the robustness measures using singular values in this examples when compared with those of the previous example.

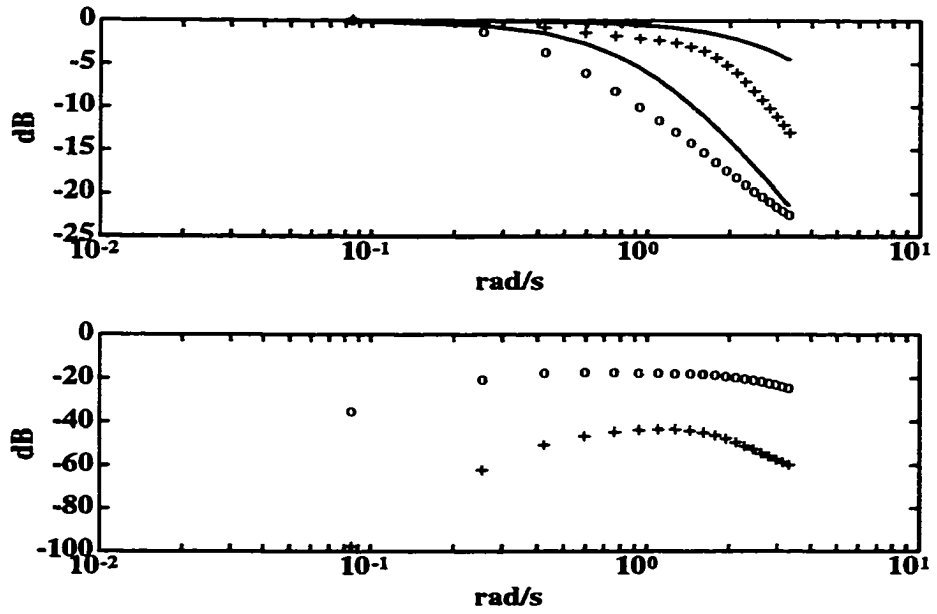


Figure 3.5: Desired and achievable closed loop frequency response for MIMO example with practical constraints: diagonal elements (top) and off-diagonal elements (bottom)

3.5 Conclusion

In this chapter the SISO PID controller tuning using finite, discrete frequency response data has been extended to the MIMO case. A stability constraint and stability margin checking or penalization through singular values are the penalty terms replacing the SISO penalty functions on gain and phase margins. Examples indicating the successful tuning possible for MIMO systems using this technique have been shown in this chapter.

Chapter 4

MIMO Experiments

In this chapter, the PID controller tuning technique developed in the last chapter is applied to an experimental system in the laboratory. The experimental study is the first attempt at applying this technique to a real-world problem as opposed to a "nice" conveniently available transfer function as was done in the examples of the previous chapter. At the point of writing this chapter, the experimental study is not thorough, however it indicates the potential of the method.

After describing the experimental apparatus, preliminary trials of identification of the system at two extremes of operation, tuning is performed and the results discussed.

4.1 Experimental Apparatus

The PID tuning for MIMO systems using finite frequency response technique is tested on a distillation column in the Chemical Engineering unit operations laboratory at Lakehead University. The distillation column is a 10 cm diameter, 11 tray, Q.V.F.^{TM} pyrex column supplied by Pegasus Ltd. of Agincourt, Ontario. The column usually runs on methanol-water. The bottoms composition is regulated through temperature control at tray 2, cascaded to the reboiler steam flow control. The temperature of the top tray (tray 11) is controlled by varying the reflux ratio. The feed flow is controlled to prevent any upset in the feed through out the experiment. A schematic is given in Figure 4.1 below. The control equipment consist of Bailey Net90^{TM} hardware, configured to provide accessible 4-20 mA signals for setpoints, process measurements, and controller outputs. Data acquisition was done through an RTI815F board with a 12 bit AtoD converter from Analog Devices Inc. mounted in a DOLCH^{TM} portable 486DX2/66MHz computer. Some analog preprocessing was required to prevent information loss at the AtoD conversion. The 4-20 mA currents were converted to voltages using 50Ω resistors, giving a range of 0.2-1 V as the available signal range corresponding to 0-100 C span. For the typical 92-94 C variations of setpoint, the corresponding voltage change across the resistor was typically 0.016V on an average of 0.944V. If the voltage across the resistor were sampled directly then this 16mV signal of interest would affect at most the bottom 5 bits of the AtoD converter (bipolar range). In

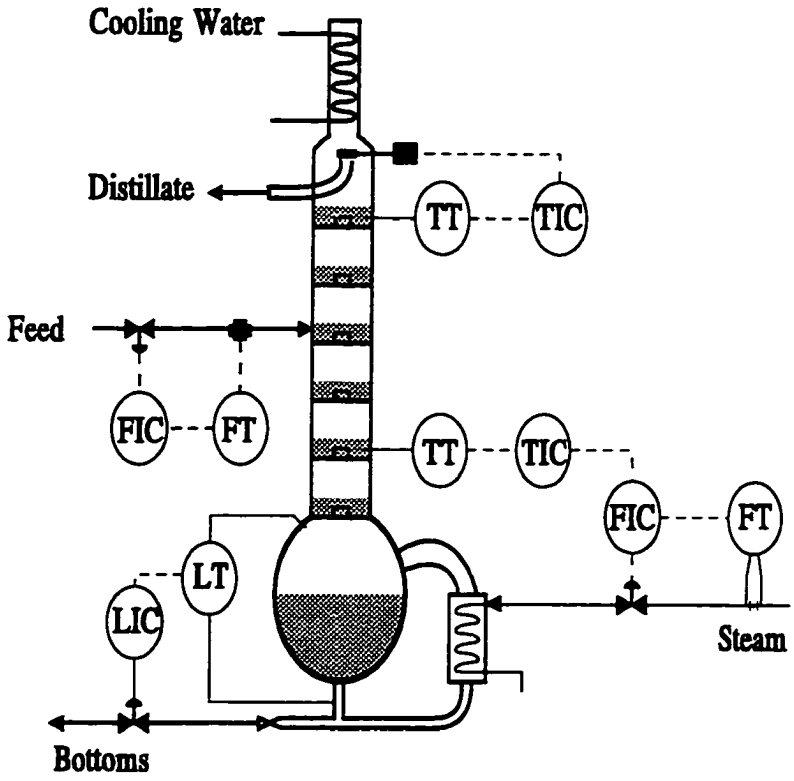


Figure 4.1: Distillation column schematic diagram

addition, AtoD quantization noise usually reduces the available signal by another bit. No amount of digital processing thereafter can improve this poor resolution. The situation calls for analog preprocessing of the signals by an instrumentation amplifier to convert the differential signal across the resistor in the 4-20 mA loop into a grounded voltage, followed by a stage of bias removal and amplification of the remaining 16 mV to a desired level before quantizing the same through the AtoD converter.

4.2 Closed Loop Band-Pass Filter Identification

The distillation column described above is modeled as a 2-input 2-output system. The closed loop transfer function describing the system can be written as

$$\begin{bmatrix} y_{T_{bottom}} \\ y_{T_{top}} \end{bmatrix} = \begin{bmatrix} T_{11} & T_{12} \\ T_{21} & T_{22} \end{bmatrix} \begin{bmatrix} r_{T_{bottom}} \\ r_{T_{top}} \end{bmatrix} \tag{4.1}$$

Where $r_{T_{bottom}}$, $r_{T_{top}}$ are setpoints for lower tray temperature and upper tray temperature respectively, and $y_{T_{bottom}}, y_{T_{top}}$ are the lower tray temperature and upper tray temperature respectively. The BPF identification method (Chapter 1) was used to identify the preliminary closed loop finite frequency response data for two cases of low and high feed methanol concentrations. This is done to get a feel for the MIMO performance tuning

to be done subsequently. The top and bottom temperature of the distillation column were controlled with individual SISO PID controllers. The sampling period for this experiment was 3s and the excitation was periodic temperature setpoint bumps between 92-94°C for bottom temperature loop and 64-66°C for top temperature loop, at an approximate period of 1800s for four cycles using a stop watch. The bottom temperature setpoint and output, $r_{T_{bottom}}$ and $y_{T_{bottom}}$ respectively, and the top temperature setpoint and output, $r_{T_{top}}$ and $y_{T_{top}}$ respectively, were sampled for identification of all the finite frequency response elements of Equation (4.1) In the first step of identification, $r_{T_{bottom}}$ was excited with $r_{T_{top}}$ being constant. In the second step, $r_{T_{top}}$ was excited with $r_{T_{bottom}}$ being constant. In both steps all control loops are in closed loop operation. The sampled data from the first step were used to identify the finite frequency response points for T_{11} and T_{21}. Similarly the sampled data from the second step were used to identify the finite frequency response points for T_{22} and T_{12}. The total operational time to identify the finite frequency response of all elements of the transfer function matrix T is $4 \times 1800 \times 2s$ (4 hours). From the closed loop frequency response matrix $T(j\omega_i)$, estimated at frequencies $\omega_i = 2\pi f_{c_i}$ and with diagonal PID controller $C(j\omega_i)$ the open loop process frequency response characteristics from the tests can be found as:

$$G(j\omega_i) = T(j\omega_i)(I - T(j\omega_i))^{-1}(C(j\omega_i))^{-1} \tag{4.2}$$

In Figures 4.2, 4.3, a typical plot of the setpoint changes and the output signals of the process for both top temperature and bottom temperature are shown, as sampled by the computer after analog preprocessing discussed in Section 4.1. The noise on the setpoints is inherent to our Bailey hardware and is unexpected of setpoint changes. The process output response for both top and bottom temperature in Figures 4.2, 4.3 is clearly nonlinear since the response in the up swing is different from the response in the down swing. For presentation purposes in these Figures the off-diagonal response to the setpoint change considered is shown offset by 0.2V. For the sake of preliminary analysis using transfer functions the open loop transfer function matrix of the distillation column can be postulated as

$$G = G_{coupling} G_{diagonal} \tag{4.3}$$

Where $G_{diagonal}$ given in Equation (4.5) is a diagonal transfer function matrix representing the open loop transfer function for each main loop. $G_{coupling}$given in Equation (4.4) is a transfer function matrix representing the coupling between loops due to interaction.

$$G_{coupling} = \begin{bmatrix} 1 & G_{coupling_{12}} \\ G_{coupling_{21}} & 1 \end{bmatrix} \tag{4.4}$$

$$G_{diagonal} = \begin{bmatrix} G_{diagonal_{11}} & 0 \\ 0 & G_{diagonal_{22}} \end{bmatrix} \tag{4.5}$$

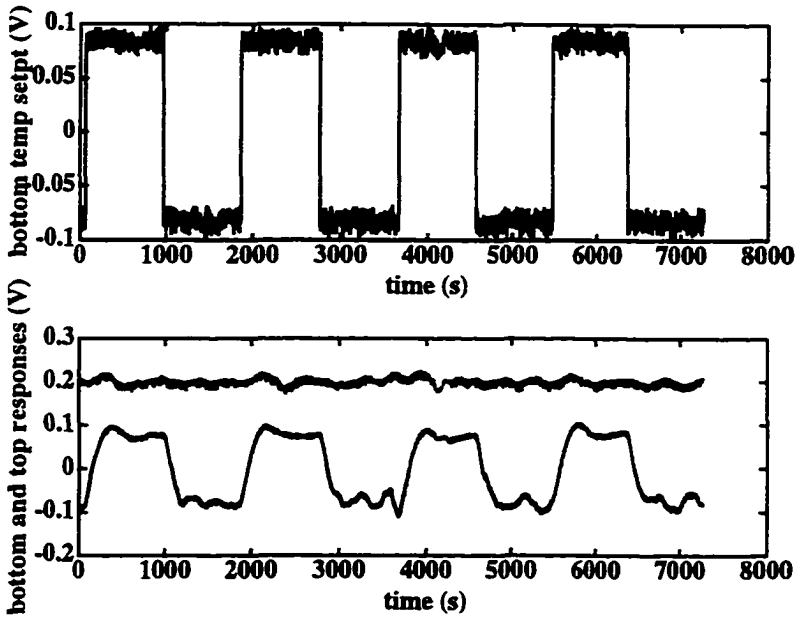


Figure 4.2: Typical change in $r_{T_{bottom}}$ and response of $y_{T_{bottom}}$ and $y_{T_{top}}$ with SISO PID controllers on column

4.2.1 Distillation Column Identification with Low Methanol Concentration

The finite frequency response for each element in the matrix G are shown in Figures 4.4, 4.5 A transfer function is fitted in to the finite frequency response data for each element of G by trial and error assuming that the coupling elements are gain+delay. The fitted open loop transfer function is given in Equation (4.6) below.

$$G(s) = \begin{bmatrix} 1 & 0.47e^{-130s} \\ 0.16e^{-110s} & 1 \end{bmatrix} \begin{bmatrix} \frac{0.0009e^{-30s}}{1+55.5s} & 0 \\ 0 & \frac{0.0016}{1+40} \end{bmatrix} \tag{4.6}$$

4.2.2 Distillation Column Identification with High Methanol Concentration

The finite frequency response for each element in the matrix G are shown in Figures 4.6,4.7.

A transfer function is fitted to the finite frequency response data for each element of G by trial and error assuming that the coupling elements are gain+delay. The fitted open loop transfer function is given in Equation (4.7)

$$G(s) = \begin{bmatrix} 1 & 1.0e^{-20s} \\ 0.1334e^{-115s} & 1 \end{bmatrix} \begin{bmatrix} \frac{0.0015e^{-5s}}{1+40s} & 0 \\ 0 & \frac{0.0025}{1111.1s^2+47.13s+1} \end{bmatrix} \tag{4.7}$$

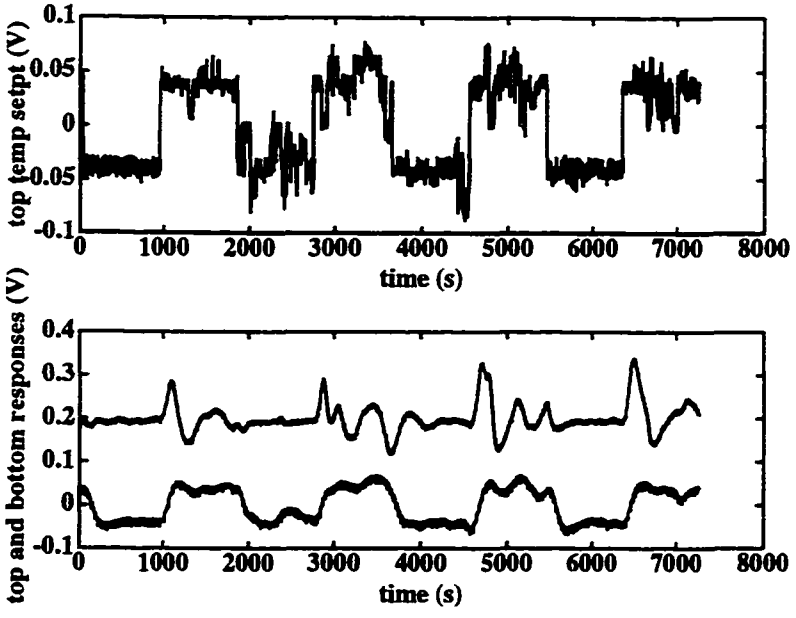


Figure 4.3: Typical change in $r_{T_{top}}$ and response of $y_{T_{top}}$ and $y_{T_{bottom}}$ with SISO PID controllers on column

Note in the case of high methanol concentration the DC gains of the transfer function matrix has increased which is to be expected. It is not yet quite clear why the delays in $G_{coupling}$ between top and bottom temperatures are very different in this case.

4.3 DC Analysis of Controller Tuning

From the preliminary transfer function results obtained in Equation (4.6) and Equation (4.7), it was observed that the modified Nyquist plot (Section 3.3) crosses the real axis at DC frequency close to the -1 point for the PID gains considered in the identification experiment (where the decoupling matrix K is an identity matrix). To apply the optimization technique a wedge with a vertex on the negative real axis has to be imposed to ensure that the modified Nyquist plot during the optimization does not encircle the -1 point. The question is how to choose the vertex for this wedge. To see if it is realistic to place the vertex at say -0.5 (corresponding $\frac{1}{G_m}$ of 6dB), we can do the following DC gain analysis. The DC gain of the modified$|(I + GKC|) - 1$ (with $s = 0$) is given by Equation (4.8).

$$|G(0)||K|k_{I_1}k_{I_2} - 1 \tag{4.8}$$

In Equation (4.8), k_{I_1},k_{I_2} are the integral gains of the PID controllers on the first and the second loop respectively. One can now estimate the required $|K|$ to achieve a vertex of -0.5 for the modified$|(I + GKC)| - 1$. The required $|K|$ to move the DC point to a vertex of -0.5 for both low and high methanol concentration are 12.5E3 and 6E3 respectively. It is impossible in both cases of low and high methanol concentrations to achieve such a

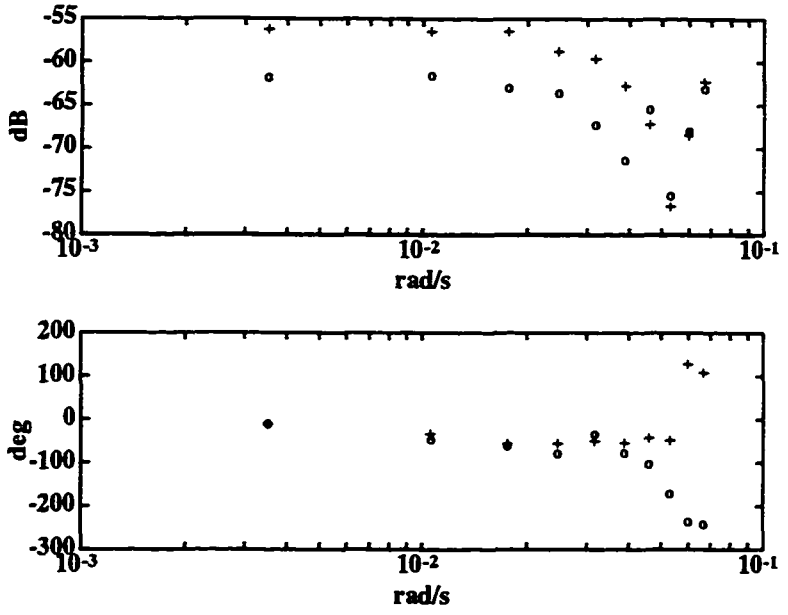


Figure 4.4: Finite frequency response data of process G with low methanol concentration in feed — g_{11} = 'o' g_{22} = '+'

desired vertex as the decoupling gain are high and other frequency response points (other than DC) encircle the -1 point of the modified$|(I+GKC)|-1$. For this reason the vertex chosen here for stability assessment was 0.99. The angle of the wedge was 2° (1 on each side).

4.4 Identification and Tuning of the Distillation Column

The PID tuning technique using finite frequency response data is implemented on the distillation column described above, after identifying it again on a different day with unknown methanol concentrations. The closed loop frequency response points are identified using the BPF method. The open loop frequency response data are calculated using Equation (4.2). The open loop frequency response data for the diagonal and off-diagonal elements of G are shown in Figures 4.8,4.9.

For tuning, the system is assumed to be excited at its set points by periodic trapezoidal waveforms with a period of 1800s, amplitude of 1 and slope of 1/10. 10 odd harmonics of the input excitation and process frequency response are used for minimizing the function in Equation (3.25). The singular value penalty function (Equation (3.33)) is used in the tuning with c_1 and c_2 values of 0.7. The desired transfer function matrix is given in Equation (4.9) below. The dead time τ_d is estimated using the open loop finite frequency response data. The natural frequency ω_n is estimated using the closed loop finite frequency

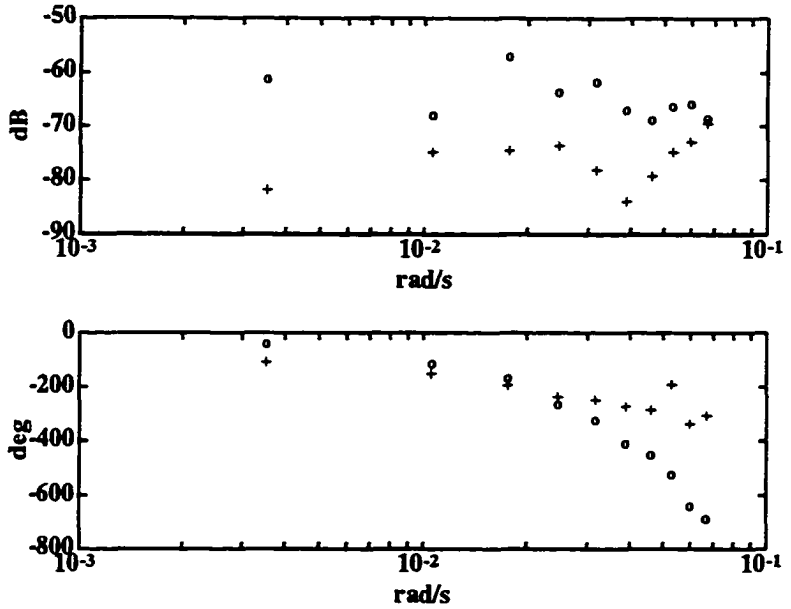


Figure 4.5: Finite frequency response data of process G with low methanol concentration in feed — g_{12} = 'o' g_{21} = '+'

response points. The damping factor for both loops is selected as 0.9 rad/s to decrease the over shoot.

$$T_d(s) = \begin{bmatrix} \frac{0.0001e^{-15s}}{s^2+0.018s+0.0001} & 0 \\ 0 & \frac{0.0001}{s^2+0.018s+0.0001} \end{bmatrix} \tag{4.9}$$

The Nelder-Mead optimization technique is used to arrive at the results given in Equations (4.10, 4.11)

$$C(s) = \begin{bmatrix} 40 + \frac{3.66}{s} & 0 \\ 0 & 40 + \frac{3.6}{s} \end{bmatrix} \tag{4.10}$$

$$K = \begin{bmatrix} 1 & -0.63 \\ 0 & 1 \end{bmatrix} \tag{4.11}$$

The results given in Equations (4.10, 4.11) were implemented on the distillation column for four hours, after which bumps were applied to record the time domain responses shown in Figures 4.10, 4.11. In these Figures, the off-diagonal responses are offest by 0.2V for presentation purposes.

4.4.1 Discussion

In the tuning of this experiment the off-diagonal terms in Equation (3.25) were weighted with a weight of 1 relative to main diagonal errors. As a result coupling still exists between the responses. If in a new tuning, the weighting on the off-diagonal elements of Equation (3.25) are increased by a significant amount the coupling can be reduced

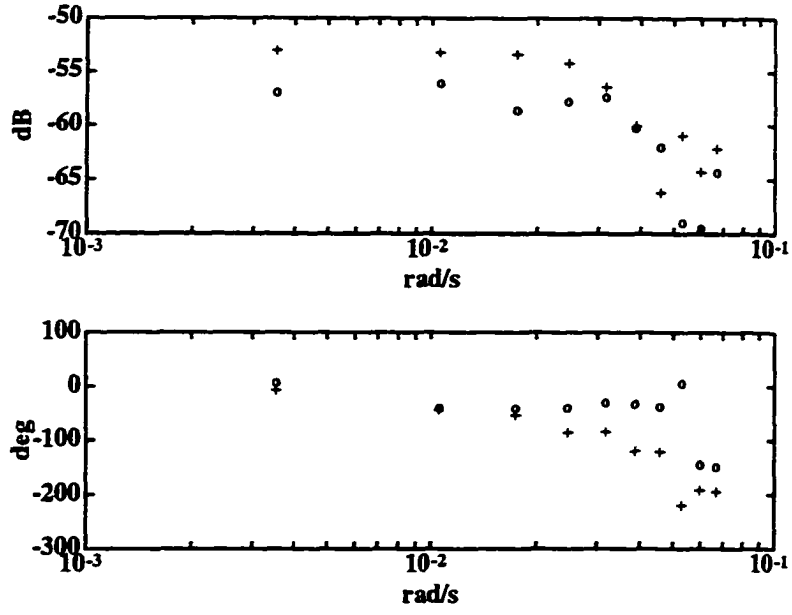


Figure 4.6: Finite frequency response data of process G with high methanol concentration in feed — g_{11} = 'o' g_{22} = '+'

further. In Figure 4.11 the time domain response in the second bump is unsatisfactory due to flooding of the trays in the distillation column while recording the data. Such flooding is usually uncommon behaviour in such systems and when it occurs usually clears on its own.

Note that the gains K recommended in Equation (4.11) are $k_{12} = -0.63$ to decouple the top to bottom response coupling while $k_{21} = 0$ indicating that the bottom to top coupling is pretty much decoupled as is. These results are borne out when comparing Figures 4.2, 4.3 with Figures 4.10, 4.11.

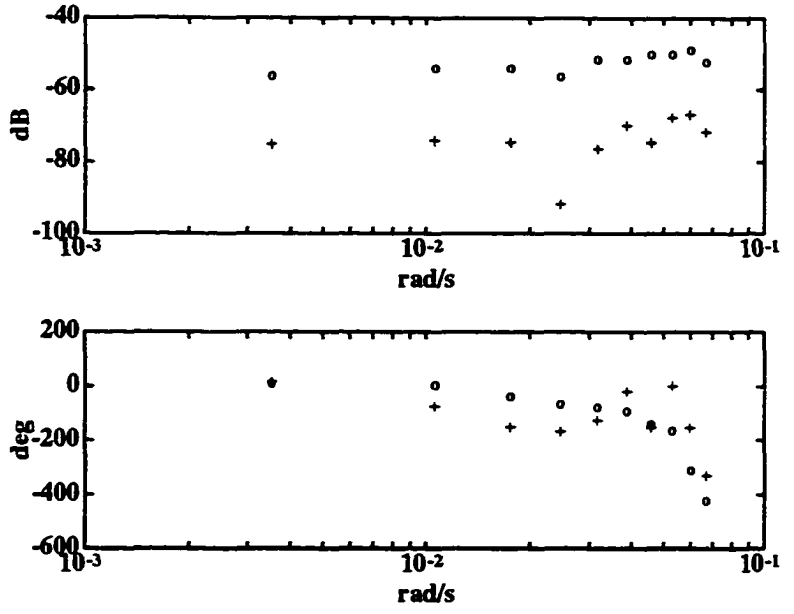


Figure 4.7: Finite frequency response data of process G with high methanol concentration in feed — g_{12} = 'o' g_{21} = '+'

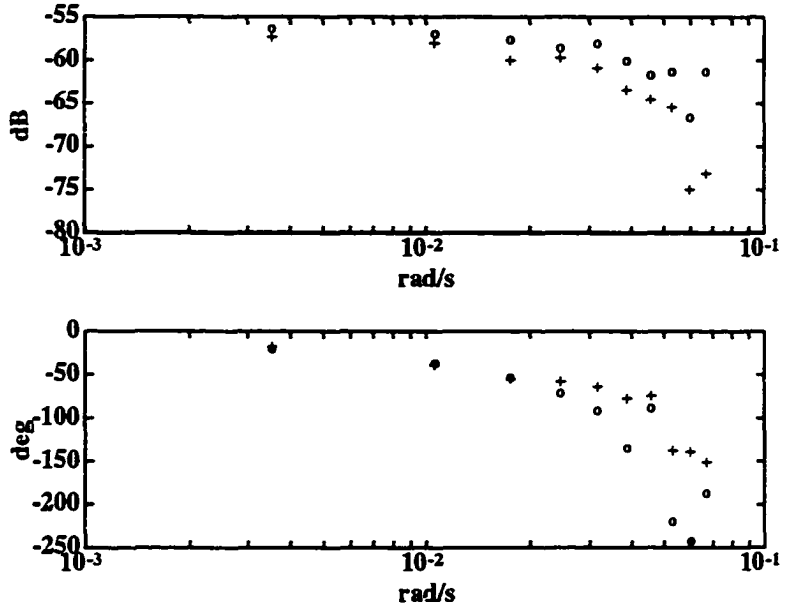


Figure 4.8: Finite frequency response data of process G on which tuning is performed — g_{11} = 'o' g_{22} = '+'

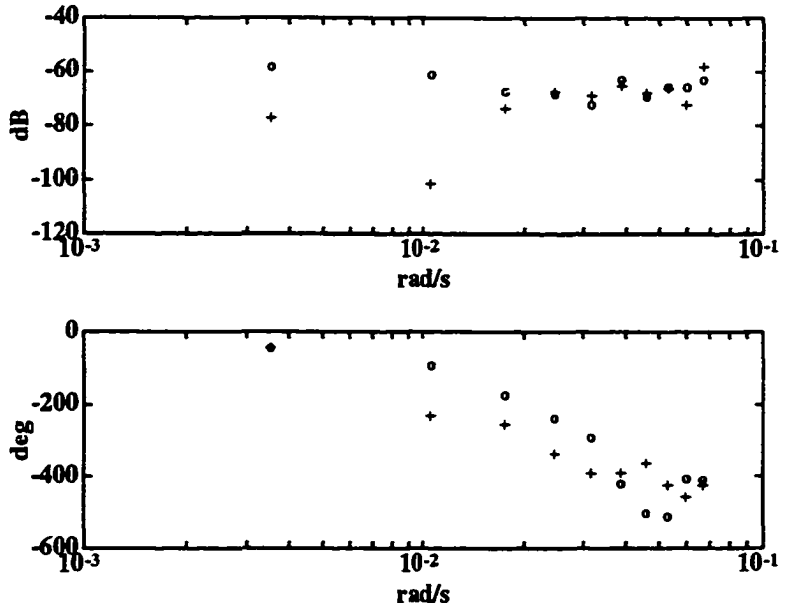


Figure 4.9: Finite frequency response data of process G on which tuning is performed — g_{12} = 'o' g_{21} = '+'

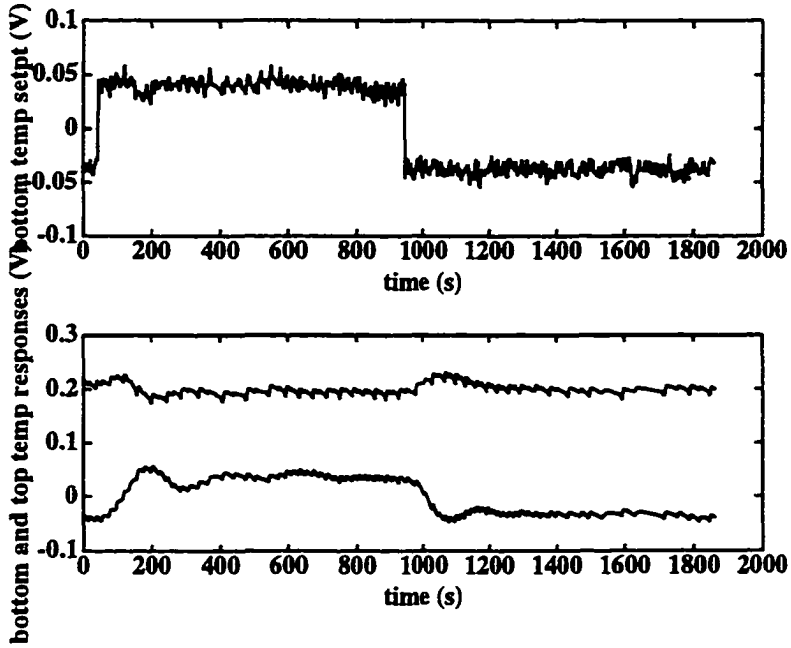


Figure 4.10: Tuned column response to change in $r_{T_{bottom}}$ with MIMO PID controller on column

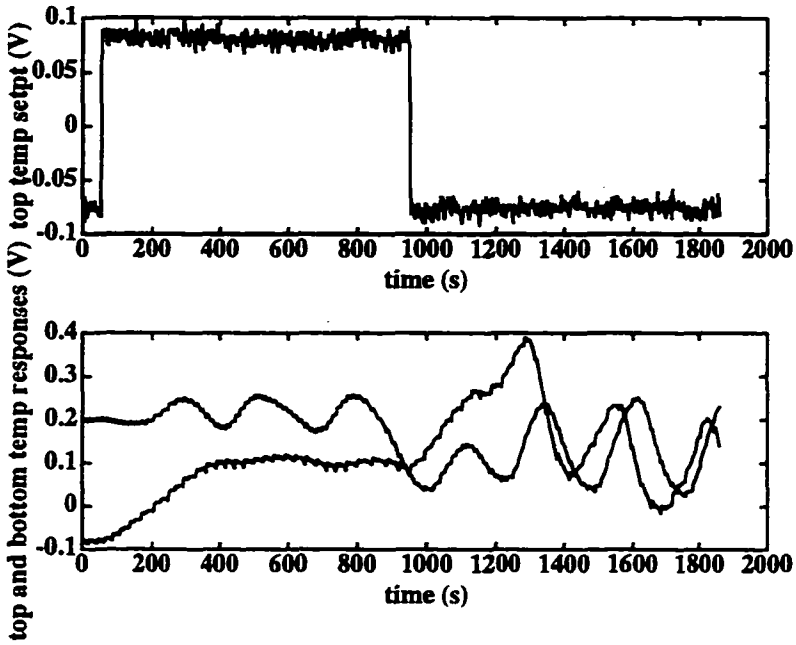


Figure 4.11: Tuned column response to change in $r_{T_{top}}$ with MIMO PID controller on column

Chapter 5

Suggestions for Future Work

More experimental work needs to be done on the distillation column for the new MIMO PID technique using finite frequency response points. The off-diagonal elements of Equation (3.22) need to be weighted by a significant amount relative to the main diagonal errors to reduce coupling between the loops.

The optimization technique used in this work for MIMO systems is sensitive to the initial conditions and therefore needs improvement. A way out may be by the use of for example simulated annealing [35].

To use this method in industry, it is required that the work be ported to an industrial DCS. The SISO work has been ported to a Bailey DCS but the MIMO has yet to be done.

The work also shows potential for extension to adaptive control in the frequency domain.

Bibliography

[1] Young P., "*Recursive Estimation and Time-Series Analysis* ", Springer Verlag, New York, 1984

[2] Oppenheim, A.V. & Schaeffer, R.W. "*Digital Signal Processing*", Prentice Hall, 1974

[3] Niu, S. & Fisher, D.G. "Simultaneous Structure Identification and Parameter Estimation of Multivariable Systems", Int. Jou. of Control, Vol. 59, pp. 1127-1141, 1994

[4] Middleton, R.H., "Frequency Domain Adaptive Control", IFAC workshop on Robust Adaptive Control, Australia, 200–205, 1988

[5] Goberdhansingh, E., Wang, L. & Cluett, W.R., "Robust Frequency Domain Identification", Chem. Engg. Sci., Vol. 47, 1989-1999, 1989

[6] Wang, L. & Cluett, W.R., "Optimal choice of Time-Scaling Factor for Linear System Approximations Using Laguerre Models", IEEE Trans. Automatic Control, Vol. 39, pp. 1463-1467, 1994.

[7] Zervos, C.C. & Dumont, G.A., "Deterministic Adaptive Control Based on Laguerre Series Representation", Int. Jou. of Control, Vol. 48, pp. 2333-2339, 1988.

[8] LaMaire, R.O., Valavani, L., Athans, M. & Stein, G., "A Frequency Domain Estimator for Use in Adaptive Control", Automatica, 27, 23-28, 1991.

[9] Fu, Y. & Dumont, G.A., "Optimum Time Scale for Discrete Laguerre Network", IEEE Trans. Automatic Control, Vol. 38, pp. 934-938, 1993.

[10] Arfin, N., Wang., L., Goberdhansingh, E. & Cluett, W.R., "Identification of the Shell Distillation Column Using the Frequency Sampling Filter Model", Process Control, Vol. 5, 1995

[11] Wang, L. & Cluett, W.R., "A More Efficient Way to Estimate Step Response Coefficients Using the FSF Model", Proc. American Control Conference, Baltimore, 1994.

[12] Wang, L. & Cluett, W.R., "Building Transfer Function Models from Noisy Step Response Data Using the Laguerre Network", Chem. Eng. Sci., Vol. 50, 1995.

[13] Barnes, T.J.D., Wang, L., & Cluett, W.R., "A Frequency Domain Design Method for PID Controllers", Proc. of American Control Conference, San Francisco, CA., 890–894, 1993.

[14] Wang, L. & Cluett, W.R., "Estimation of Process Step Response Weights Based on Closed-loop Step Response Data", IEE Proc. Pt. D Control Theory and Applications. Vol. 141, 1994.

[15] Natarajan, K. & Gilbert, A.F., "System Identification and Controller Tuning using Bandpass Filters", To appear in Can. Jou. of Chem. Engg., Aug. 1997.

[16] Natarajan, K. and Gilbert, A.F., "Direct PID Controller Tuning Based on Finite Frequency Response Data", To appear in ISA Trans.

[17] Rudin, W. "*Real and Complex Analysis*", McGraw Hill Inc., 1987.

[18] Nelder, J.A. and Mead, R., "A Simplex Method for Function Minimization", Computer Jou., 7, 308–313, 1965.

[19] Ziegler, J.G. & Nichols, N.B., "Optimal Settings for Automatic Controller", Trans. ASME, 64, 759, 1942.

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

[21] Natarajan, K., Gilbert, A.F., Yousef, A., "PID Controller Tuning Based on Finite Frequency Data", CPPA. Proc. Control Systems., Halifax, 1996. To appear in Pulp and Paper Canada.

[22] RosenBrock, H.H., "*Computer-Aided Control System Design*", Academic Press, 1974.

[23] McAvoy, T.J., "Interaction Analysis Principles and Applications", Instrument Society of America, Monograph series 6, 1940.

[24] Munro, N., "Modern Approaches to Control System Design", Institute of Electrical Engineers. IEE control engineering series; 9, 1979.

[25] Shinskey, F.G., "Process Control Systems: Application, Design, and Tuning", McGraw-Hill, inc. Third Edition, 1979.

[26] Lehtomaki, N.A., Sandell, N.R., JR., and Athans, M., "Robustness Results in Linear-Quadratic Gaussian Based Multivariable Control Designs", IEEE Trans. AC-26, 1981.

[27] Van de Vegte, J., "*Feedback Control Systems*", Second Edition, Prentice Hall, 1990.

[28] Astrom, K.J. and Wittenmark, B. "*Adaptive Control*", second edition, Addison Wesley Publishing ,1995.

[29] B.R Barmish "*New Tools for Robustness of Linear Systems*", Macmillan, 1994 Publishing, 1994

[30] Francis, .B .,A "*A course in H_∞ control theory*", Springer-Verlag, 1987

[31] Doyle, J.C., Francis, B.A., Tannenboum, A.R., "*Feedback Control Theory*", Macmillan publishing, 1992.

[32] Youla, D.C., Jabr H.A., Bongiorno J.J, "Modern Wiener-Hopf Design Of Optimal Controllers (Part I & Part II), IEEE.AC, 1976

[33] Pulp and Paper Manufacturing Series, "*Mill-Wide Process Control and Information Systems*", Vol. 10, The joint Textbook committee of the Canadian pulp and paper association, 1993.

[34] Entech Corp. "seminar notes" provided by Dr. A.F. Gilbert

[35] W.H. Press, S.A. Teukolsky, W.T. Vetterling, B.P. Flannery, "*Numerical Recipes in C*", Cambridge University Press, 1994

VITA

Candidate's full name:	Yousef, Ali
Universities attended:	Lakehead University, Thunder Bay Ontario, Canada
Periods	1989-1993 B.Eng. (EE)
	1995-1997 Graduate Student

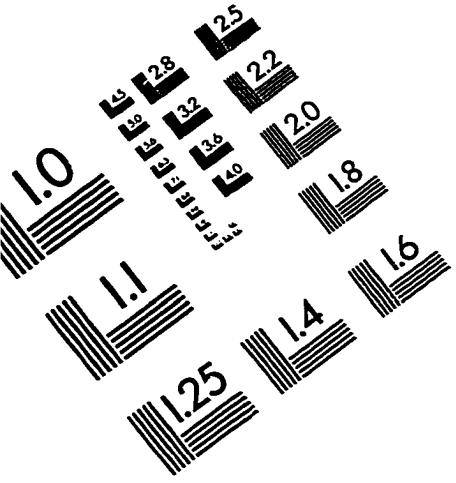
Publications:

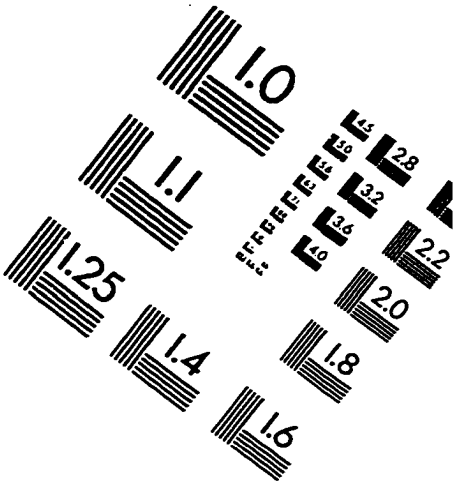
1. Natarajan, K., Gilbert, A.F., & Yousef, A., "PID Controller Tuning Based on Finite Frequency Data", presented at the C.P.P.A. Control Systems Conference in Halifax 1996. (To appear in Pulp and Paper Canada).

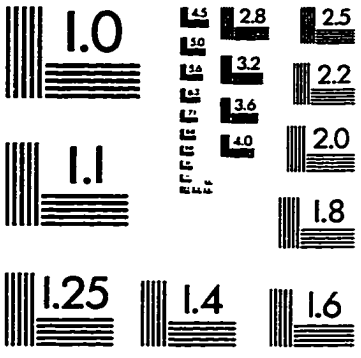
2. Natarajan,K., Yousef, A. & Gilbert A.F. "A Comparative Study of Three Closed Loop System Identfication Techniques", (Manuscript in preparation – Chapter 1 of this Thesis).

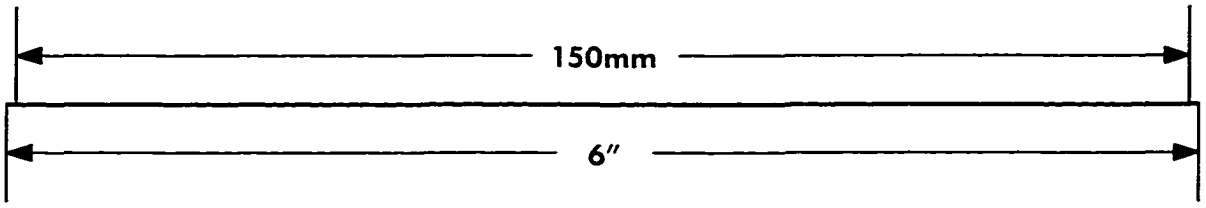
3. Yousef, A., Natarajan, K. & Gilbert, A.F. "PID Controller Tuning For MIMO systems using Finite Frequency Response", (Manuscript in preparation – Chapter 3 and Chapter 4 of this Thesis).

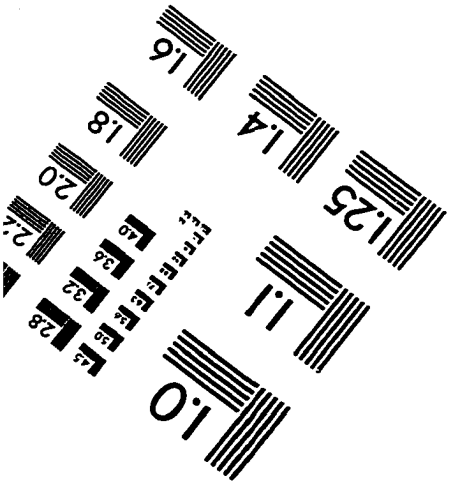
IMAGE EVALUATION
TEST TARGET (QA-3)

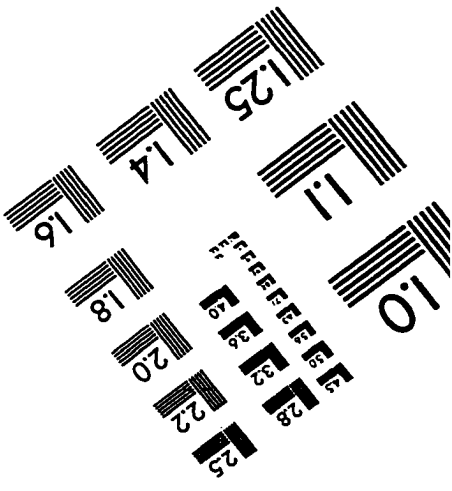












APPLIED IMAGE, Inc
1653 East Main Street
Rochester, NY 14609 USA
Phone: 716/482-0300
Fax: 716/288-5989