ISSN 0280–5316
ISRN LUTFD2/TFRT--5592--SE

Automatic Tuning of Multivariable PID Controllers

Valon Recica

Department of Automatic Control
Lund Institute of Technology
February 1998

Department of Automatic Control
Lund Institute of Technology
Box 118
S-221 00 Lund Sweden

Document name
MASTER THESIS

Date of issue
February 1998

Document Number
ISRN LUTFD2/TFRT--5592--SE

Author(s)
Valon Recica

Supervisor
Karl Henrik Johansson

Sponsoring organisation

Title and subtitle
Automatic Tuning of Multivariable PID Controllers

Abstract

This Master's thesis project is part of my education within the Mat-Nat program at Uppsala University. I was given the opportunity to perform it as an exchange student at the Department of Automatic Control in Lund.

In this Master's thesis project, MIMO control was studied on a coupled process that consists of four interconnected water tanks. The target was to control the levels in two of the tanks with two pumps.

Various methods for tuning MIMO controllers based on relay identification experiments were studied. These methods are extensions to the SISO relay tuning technique developed by Åström and Hägglund. The main idea is to identify the critical gain and the critical frequency of the process, and then use this information to tune PID controllers by the Ziegler-Nichols modifications.

The linearized dynamics of the tank process contain a multivariable zero, whose location can be easily varied to give a system that is more or less easy to control. This feature was used in the project to investigate MIMO control structures such as diagonal control and full MIMO control.

The tuning method was implemented on a PC and a man-machine interface was developed.

Key words
automatic tuning; multivariable control; decentralized PID control; multivariable zeros; decentralized relay feedback

Classification system and/or index terms (if any)

Supplementary bibliographical information

ISSN and key title
0280–5316

ISBN

Language
English

Number of pages
52

Recipient's notes

Security classification

The report may be ordered from the Department of Automatic Control or borrowed through:
University Library 2, Box 3, S-221 00 Lund, Sweden
Fax +46 46 222 44 22 E-mail ub2@ub2.lu.se

Contents

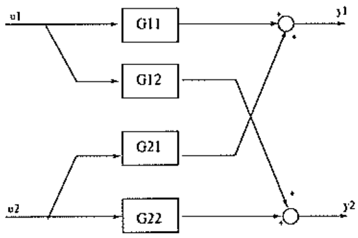

Figure 1 A system with two inputs and two outputs.

1. Introduction

Many systems in industry are multivariable, meaning that they have more than one input and more than one output. Additionally there exist interactions between the different inputs and outputs in such way that one input signal has influence on more than one output signal. A system with two inputs and two outputs can be described as shown in Figure 1. In Figure 1 we call G_{11} and G_{22} the main transfer functions, while G_{12} and G_{21} are called interactions elements.

A control system for controlling y_1 and y_2 can be described as shown in Figure 2. As for single-input-single-output systems there are some system criteria such as stability, speed and static accuracy that must be satisfied. Another important criteria is low interaction, which means that we can change the value of one of the outputs without large influence on the other output.

Poorly tuned control loops represent a large economic cost for industry. Control parameters are often manually tuned. Most modern multivariable control design methods require a full model of the system. In many cases such a model is not available. Then we need to identify the system which usually takes time and requires very good engineering skills. For SISO systems there exist simple methods for automatic tuning of SISO control loops [Åström and Hägglund, 1984].

In this master's thesis, we try to extend the SISO tuning technique to multivariable systems. It will be applied to the multivariable laboratory system, consisting of four interconnected water tanks. The target is to control the level in the bottom two tanks with the help of two pumps. This laboratory system will be referred in the following sections as the quad tank system.

The linearized model of the quad tank system has a multivariable zero which can be located in either the left or the right half plane by simply changing the positions of the valves, γ_1 and γ_2. See Figure 3. The influence of the location of the multivariable zero on control performance will be also investigated. This process has been previously studied in [Nunes, 1997] and [Johansson, 1997]. We will study the control results achieved by extending the methods used in SISO systems to our multivariable system. This method will be referred to as "SISO Controller Tuning".

An extension of the single-loop relay auto-tuner to the MIMO case will be studied. This method was proposed by Vasnani [Vasnani, 1994] and is a combination of sequential loop closing and single-loop relay tuning. To tune each loop a single relay is used to determine the corresponding critical point, and the Ziegler-Nichols settings with some modifications are then employed, [Ziegler and Nichols, 1943]. This method will be referred to as "Sequential Controller Tuning".

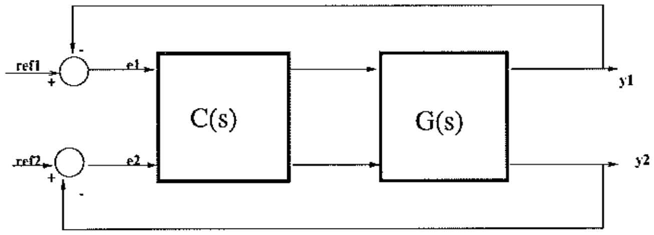

Figure 2 A multivariable control system.

A modification of the Sequential method is also discussed, [Johansson, 1997]. In this method we try to identify a more complex model of the system and thereby design a better controller. This method will be referred to as "Extension to Sequential Controller Tuning".

A complete new method for auto-tuning fully cross-coupled multivariable PID controllers from decentralized relay feedback will be also studied, [Wang et al., 1997]. First, from the decentralized test, the frequency response-matrix and steady-state matrix are identified. Then, a new set of design equations are used to design the controller parameters. This method will be referred to as "Decentralized Controller Tuning".

The outline of the thesis is as follows: In Section 2 we describe the quad tank system and refer to its main characteristics. We derive also the physical transfer functions from control signals to measurement signals. In this Section we discuss also the pairing problem, i.e., how the pairing of the manipulated and controlled variables influences the ability to control the system. The introductory part of Section 3 contains review material about automatic tuning, which helps the reader to develop an understanding of what automatic tuning is about. Thereafter we perform the experiments associated with the methods described above. The experiments are done using two different configurations of the system, a minimum phase system, where both the systems zeros are in the LHP, and a non minimum phase system, where one of the systems zeros is in the RHP. In addition, for the non minimum phase system we perform experiments for two various control structures, i.e., the manipulated and control variables are paired in two possible ways. In Section 4, the design of a user friendly interface is described. The software used for developing the graphical interface was InTouch, which is a software used to create PC based man-machine interfaces.

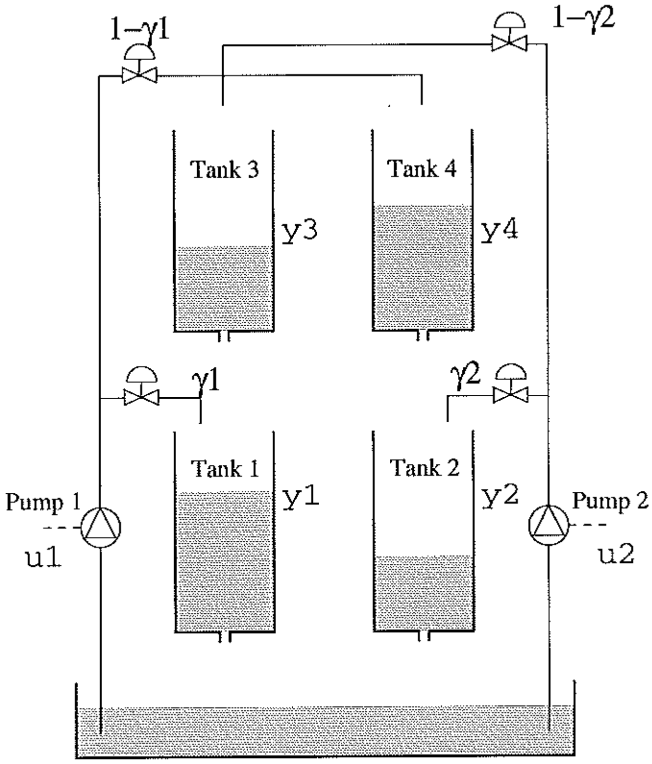

Figure 3 The Quad-Tank process.

2. The Quad Tank System

In this Section we derive a mathematical model for the quad tank system, shown in Figure 3, from physical data. The quad tank system is a non-linear system. It consists of four interconnected water tanks. The inputs of the system are u_1 and u_2 (the input voltages to the pumps) and the outputs of the system are y_1 and y_2 (the voltages from the measurement devices). There are also two valves γ_1 and γ_2, which influence the balance of the inflows between the upper and lower tanks. It will be shown later that, by changing the position of the valves, we also change the properties of the system, i.e., the location of one of the system zeros can be easily varied to give a system which is more or less easy to control.

First, we will derive the nonlinear differential equations and then linearize them around a chosen operating point.

2.1 Derivation and linearization of physical model

A tank with inflow q_{in} and outflow q_{out} is described with the following relations (the mass balance law):

$$\frac{dV}{dt} = q_{in} - q_{out} \tag{1}$$

where V is the volume of the tank
q_{in} is the inflow
q_{out} is the outflow

Let A be the cross-section area and h the water level. Then

$$\frac{dV}{dt} = A\frac{dh}{dt} \tag{2}$$

Equations (1)–(2) give

$$\frac{dh}{dt} = \frac{q_{in}}{A} - \frac{q_{out}}{A} \tag{3}$$

The Bernoulli's law gives us

$$q_{out} = a\sqrt{2gh} \tag{4}$$

We assume that the flow generated by the pump is proportional to the applied voltage, i.e.,

$$q_{pump} = q_{in} = ku \tag{5}$$

This is motivated by the fact that the time constant of the pump is small compared to the tank dynamics. In the same way, assume that the level measurement signal y is proportional to the true level h. Thus we write

$$y = k_c h \tag{6}$$

Equations (3)–(5) give the following non-linear equation

$$\frac{dh}{dt} = \frac{ku}{A} - \frac{a\sqrt{2gh}}{A}$$

Now that we have derived the non-linear equation for the single tank-process, we proceed in the same way to derive the equations for the quad tank system (see Figure 3) and after a little work we arrive at the following equations [Johansson, 1997]

$$\begin{aligned}
A_1\frac{dh_1}{dt} &= \gamma_1 k_1 u_1 - a_1\sqrt{2gh_1} + a_3\sqrt{2gh_3}\\
A_2\frac{dh_2}{dt} &= \gamma_2 k_2 u_2 - a_2\sqrt{2gh_2} + a_4\sqrt{2gh_4}\\
A_3\frac{dh_3}{dt} &= (1-\gamma_2) k_2 u_2 - a_3\sqrt{2gh_3}\\
A_4\frac{dh_4}{dt} &= (1-\gamma_1) k_1 u_1 - a_4\sqrt{2gh_4}
\end{aligned}$$

For a stationary operating point $(h_1^0, h_2^0, h_3^0, h_4^0, u_1^0, u_2^0)$, the non-linear differential equations gives that

$$\frac{(1-\gamma_2)k_2u_2^0}{A_3} = \frac{a_3\sqrt{2gh_3^0}}{A_3}$$

$$\frac{(1-\gamma_1)k_1u_1^0}{A_4} = \frac{a_4\sqrt{2gh_4^0}}{A_4}$$

and thus

$$\frac{a_1\sqrt{2gh_1^0}}{A_1} = \frac{a_3\sqrt{2gh_3^0}}{A_1} + \frac{\gamma_1 k_1 u_1^0}{A_1} = \frac{(1-\gamma_2)k_2u_2^0}{A1} + \frac{\gamma_1 k_1 u_1^0}{A_1}$$

$$\frac{a_2\sqrt{2gh_2^0}}{A_2} = \frac{a_4\sqrt{2gh_4^0}}{A_2} + \frac{\gamma_2 k_2 u_2^0}{A_2} = \frac{(1-\gamma_1)k_1u_1^0}{A_2} + \frac{\gamma_2 k_2 u_2^0}{A_2}$$

Next step is to linearize the system around a working point. Introducing $\Delta h_i = h_i - h_i^0$ and $\Delta u_i = u_i - u_i^0$ and doing Taylor expansions of the terms yields the linearized equations:

$$\begin{aligned}
A_1\frac{d\Delta h_1}{dt} &= \gamma_1 k_1 \Delta u_1 - a_1\sqrt{\frac{g}{2h_1^0}}\Delta h_1 + a_3\sqrt{\frac{g}{2h_3^0}}\Delta h_3 \\
A_2\frac{d\Delta h_2}{dt} &= \gamma_2 k_2 \Delta u_2 - a_2\sqrt{\frac{g}{2h_2^0}}\Delta h_2 + a_4\sqrt{\frac{g}{2h_4^0}}\Delta h_4 \\
A_3\frac{d\Delta h_3}{dt} &= (1-\gamma_2) k_2 \Delta u_2 - a_3\sqrt{\frac{g}{2h_3^0}}\Delta h_3 \\
A_4\frac{d\Delta h_4}{dt} &= (1-\gamma_1) k_1 \Delta u_1 - a_4\sqrt{\frac{g}{2h_4^0}}\Delta h_4
\end{aligned}$$

Define the time constants as $T_i = \frac{A_i}{a_i}\sqrt{\frac{2h_i}{g}}$ and rename the states and inputs as $x = \Delta h$ and $u = \Delta u$, respectively. Then

$$\frac{dx}{dt} = \begin{bmatrix} \frac{-1}{T_1} & 0 & \frac{A_3}{A_1T_3} & 0 \\ 0 & \frac{-1}{T_2} & 0 & \frac{A_4}{A_2T_4} \\ 0 & 0 & \frac{-1}{T_3} & 0 \\ 0 & 0 & 0 & \frac{-1}{T_4} \end{bmatrix} x + \begin{bmatrix} \frac{\gamma_1 k_1}{A_1} & 0 \\ 0 & \frac{\gamma_2 k_2}{A_2} \\ 0 & \frac{(1-\gamma_2)k_2}{A_3} \\ \frac{(1-\gamma_1)k_1}{A_4} & 0 \end{bmatrix} u$$

and

$$y = \begin{bmatrix} k_c & 0 & 0 & 0 \\ 0 & k_c & 0 & 0 \end{bmatrix} x$$

The corresponding transfer function is

$$G(s) = \begin{pmatrix} \frac{c_1\gamma_1}{1+sT_1} & \frac{c_1(1-\gamma_2)}{(1+sT_3)(1+sT_1)} \\ \frac{c_2(1-\gamma_1)}{(1+sT_4)(1+sT_2)} & \frac{c_2\gamma_2}{1+sT_2} \end{pmatrix}$$

with $c_1 = \frac{T_1 k_1 k_c}{A_1}$ and $c_2 = \frac{T_2 k_2 k_c}{A_2}$.

2.2 Multivariable Zeros

The linearized dynamics of the quad tank system contain a multivariable zero, whose location can be easily varied to give a system which is more or less easy to control. Briefly we recall the definition of zeros in SISO and MIMO systems.

The zeros of a SISO system are the roots of the numerator polynomial of its transfer function. Depending on the zero locations there are two types of systems:

Minimum phase systems All zeros are located in the left half-plane

Non-minimum phase systems At least one zero is in the right half-plane

There are many ways of defining zeros of MIMO systems. One of the definitions that can be used to define the MIMO zeros is: MIMO zeros are those values of s for which the rank of $G(s)$ drops below its nominal rank. These zeros are often called transmission zeros. In general, the presence of a MIMO zero implies a transmission blocking property, then they are invariant under feedback control, and become the poles of the inverse. These properties imply the difficulty of controlling a system with a MIMO zero.

The MIMO zeros of the quad tank system are equal to the roots of $\det G(s)$ where

$$\det G(s) = \frac{c_1 c_2}{\gamma_1 \gamma_2 \prod_{i=1}^{4}(1+sT_i)} \left[(1+sT_3)(1+sT_4) - \frac{(1-\gamma_1)(1-\gamma_2)}{\gamma_1 \gamma_2} \right]$$

$$\gamma_1, \gamma_2 \in (0,1)$$

Hence, the system has two zeros. One of these is always in the left half-plane. The location of the other zero depends however on the sign of

$$\eta = \frac{(1-\gamma_1)(1-\gamma_2)}{\gamma_1 \gamma_2}$$

It follows that the zero is in the right half-plane if $\eta < 0$, i.e., $\gamma_1 + \gamma_2 < 1$, which means that we have more flow going to the upper tanks and less flow going to the lower tanks. The zero is in the left half-plane if $\eta > 0$, i.e., $\gamma_1 + \gamma_2 > 1$, which means that we have more flow going to the lower tanks and less flow going to the upper tanks. The zero is in the origin if $\eta = 0$, i.e., $\gamma_1 + \gamma_2 = 1$, which means that the same flow is going to the upper and lower tanks.

2.3 Parameter Values

To determine the values of the physical constants in the derived model, we have to measure some components of the quad tank system (like the diameter of the holes and the tanks) and perform some simple experiments in order to find the values of the pump constants and valve positions.

The four-tank system consists of two types of tanks which have slightly different dimensions. Tank 1 and Tank 3 have dimensions

$$\begin{aligned} A &= 28 \text{ cm}^2 \\ a &= 0.071 \text{ cm}^2 \end{aligned}$$

The dimensions of Tank 2 and Tank 4 are

$$\begin{aligned} A &= 32 \text{ cm}^2 \\ a &= 0.057 \text{ cm}^2 \end{aligned}$$

The constant k_c used in the level sensors is

$$k_c = 0.50 \text{ V/cm}$$

The value of the constants associated with the pumps are

$$\begin{aligned} k_1 &= 3.1 \text{ cm}^3/\text{s} \\ k_2 &= 3.3 \text{ cm}^3/\text{s} \end{aligned}$$

The values for these constants could vary a little if the voltage fed to the pumps differs much from the ones used for the calculations.

To determine the values of γ_1 and γ_2 we need to do the following: We run the system until the steady state is reached. Thereafter, from the non-linear equations for the tanks three and four, assuming stationarity, we extract the values of γ_1 and γ_2.

2.4 Minimum Phase System

By setting the position of the valves so that we have more flow going to the lower tanks, i.e., $\gamma_1+\gamma_2 > 1$, both system zeros will be located at the LHP. After experimenting with different input voltages, we arrive at the following stationary values for the system $(h_1^0, h_2^0; u_1^0, u_2^0) \approx (12.4 \text{ cm}, 12.7 \text{ cm}; 3.0 \text{ V}, 3.0 \text{ V})$. Then using the expressions derived in this Section knowing that the stationary level of the upper tanks are $(h_3^0, h_4^0) \approx (1.8 \text{ cm}, 1.4 \text{ cm})$, we obtain the following values for the time constants, T_i, associated with each tank,

$$\begin{aligned} T_1 &= 62 \text{ s} \\ T_2 &= 90 \text{ s} \\ T_3 &= 23 \text{ s} \\ T_4 &= 30 \text{ s} \end{aligned}$$

and from the measurements we obtain

$$\begin{aligned} k_1 &= 3.33 \text{ cm}^3/\text{s} \\ k_2 &= 3.35 \text{ cm}^3/\text{s} \\ \gamma_1 &= 0.70 \\ \gamma_2 &= 0.60 \end{aligned}$$

Knowing these parameters we can derive the physical transfer function matrix of the system

$$G(s) = \begin{bmatrix} \frac{2.6}{1+62s} & \frac{1.5}{(1+23s)(1+62s)} \\ \frac{1.4}{(1+30s)(1+90s)} & \frac{2.8}{1+90s} \end{bmatrix}$$

with the transmission zeros in $z = -0.018$ and $z = -0.060$.

2.5 Non-Minimum Phase System

If we instead have less flow going to the lower tanks, i.e., $\gamma_1 + \gamma_2 < 1$, one of the zeros will be located at the RHP. Now, we try to find an operating point close to the one we got for the minimum phase setting, and after experimenting with different input voltages, we arrive at the following stationary values $(h_1^0, h_2^0; u_1^0, u_2^0) \approx (12.6 \text{ cm}, 13.0 \text{ cm}; 3.15 \text{ V}, 3.15 \text{ V})$. Then using the expressions derived in this Section knowing that the stationary level of the upper tanks are $(h_3^0, h_4^0) \approx (4.8 \text{ cm}, 4.9 \text{ cm})$, we obtain the following values for the time constants, T_i, associated with each tank,

$$\begin{aligned} T_1 &= 63 \text{ s} \\ T_2 &= 91 \text{ s} \\ T_3 &= 39 \text{ s} \\ T_4 &= 56 \text{ s} \end{aligned}$$

and from the measurements

$$\begin{aligned} k_1 &= 3.14 \text{ cm}^3/\text{s} \\ k_2 &= 3.29 \text{ cm}^3/\text{s} \\ \gamma_1 &= 0.43 \\ \gamma_2 &= 0.34 \end{aligned}$$

Then the system transfer function matrix derived by means of physical modeling is

$$G(s) = \begin{bmatrix} \frac{1.5}{1+63s} & \frac{2.5}{(1+39s)(1+63s)} \\ \frac{2.5}{(1+56s)(1+91s)} & \frac{1.6}{1+91s} \end{bmatrix}$$

with transmission zeros in $z = 0.013$ and $z = -0.057$. Note that now one of the transmission zeros is located in the right half-plane.

2.6 The Input-Output Pairing Problem

Given a system with two inputs and two outputs it is possible to choose two different input-output pairing configurations for the controllers in the corresponding multivariable interacting control systems. Hence any two-input two-output system generates at most two different systems which are more or less harder to control.

Thus, before we begin with various methods for designing controllers we need to investigate how the pairing of the measured values and control variables should be done in order to have a system which is easier to control. There is no simple method for doing this, but an indication can be obtained by the relative gain array (RGA). This can be computed from the static gains in all loops in a multivariable system. For the quad tank system the RGA is defined as

$$R = \begin{bmatrix} \lambda & 1-\lambda \\ 1-\lambda & \lambda \end{bmatrix}$$

where

$$\lambda = \frac{G_{11}(0)G_{22}(0)}{G_{11}(0)G_{22}(0) - G_{12}(0)G_{21}(0)}$$

The scalar λ has the interpretation

$$\lambda = \frac{(\frac{\partial y_1}{\partial u_1})_{open}}{(\frac{\partial y_1}{\partial u_1})_{closed}}$$

where $(\frac{\partial y_1}{\partial u_1})_{open}$ represent the steady-state gain with the first loop open and $(\frac{\partial y_1}{\partial u_1})_{closed}$ represent the steady-state gain with the second loop perfectly closed, i.e., $y_2 = 0$.

Bristol's recommendation for controller pairing is that the measured values and control variables should be paired so that the corresponding relative gains are positive and as close to one as possible [Bristol, 1966].

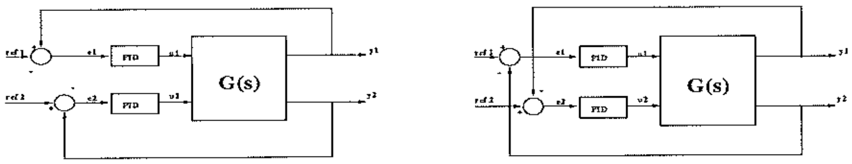

Figure 4 The Figure illustrating the pairing possibilities.

Let the system inputs and outputs be connected as in the left part of Figure 4. Then the transfer function of the system is

$$G(s) = \begin{bmatrix} G_{11} & G_{12} \\ G_{21} & G_{22} \end{bmatrix}$$

The zeros of the above system are given by

$$\det G(s) = G_{11}G_{22} - G_{12}G_{21} = 0$$

We connect the system as shown in the right part of Figure 4. Then the transfer function of the system is

$$G(s) = \begin{bmatrix} G_{12} & G_{11} \\ G_{22} & G_{21} \end{bmatrix}$$

and the zeros of the system are given by

$$\det G(s) = G_{12}G_{21} - G_{11}G_{22} = 0$$

Notice that we will have the same position of the zeros, meaning that in this case we will not have any control limitations due to the position of the zeros. Let us see how the pairing influences the minimum phase system.

Computation of λ for two different control structures:

$$\lambda_1 = \frac{G_{11}(0)G_{22}(0)}{G_{11}(0)G_{22}(0) - G_{12}(0)G_{21}(0)} = 1.4$$

$$\lambda_2 = \frac{G_{12}(0)G_{21}(0)}{G_{12}(0)G_{21}(0) - G_{11}(0)G_{22}(0)} = -0.4$$

for $G_{11}(0) = 2.6$, $G_{12}(0) = 1.5$, $G_{21}(0) = 1.4$ and $G_{22}(0) = 2.8$.

The relative gain array gives us an indication that the measured values and control variables of the minimum phase system should be paired as in the first control structure.

We do the same computation for the non minimum phase system. Computation of λ for two different control structures:

$$\lambda_1 = \frac{G_{11}(0)G_{22}(0)}{G_{11}(0)G_{22}(0) - G_{12}(0)G_{21}(0)} = -0.67$$

$$\lambda_2 = \frac{G_{12}(0)G_{21}(0)}{G_{12}(0)G_{21}(0) - G_{11}(0)G_{22}(0)} = 1.6$$

for $G_{11}(0) = 1.5$, $G_{12}(0) = 2.5$, $G_{21}(0) = 2.5$ and $G_{22}(0) = 1.6$.

The relative gain array gives us an indication that the measured values and control variables of the non minimum phase system should be paired as in the second control structure.

In the future, for the minimum phase system we will perform tuning experiments only for the plant coupled as in the first control structure. For the non minimum phase system we will perform tuning experiments for both control structures.

2.7 Comparison of Experimental Results with Simulations

In this Section we compare the results achieved from the experiments on the real system with the results achieved from the simulation of the physical linear model. The purpose of this is to show the good agreement between the results achieved from the experiments with the results achieved from the simulations. This comparison will be shown only for the non minimum phase system, control structure two, and only for the so called method "SISO Controller Tuning"(see Section 3.2).

The physical transfer function matrix of the non minimum phase system, control structure two, is

$$G(s) = \begin{bmatrix} \frac{2.5}{(1+39s)(1+63s)} & \frac{1.5}{1+63s} \\ \frac{1.6}{1+91s} & \frac{2.5}{(1+56s)(1+91s)} \end{bmatrix}$$

This transfer function matrix is used as a simulation model. For the simulated relay experiments we have used exactly the same relay parameters as we have used for the real system (see Section 3.2). The same is also valid for the design of the controller parameters, where we have used the same design equations (Section 3.2).

All simulations are performed in Simulink.

SISO Controller Tuning The results of relay experiments are shown in Figure 5 and Figure 6. In Figure 7, the responses of the simulated and real system, controlled by a decentralized PI controller (see Section 3), are given for a step change of the reference one.

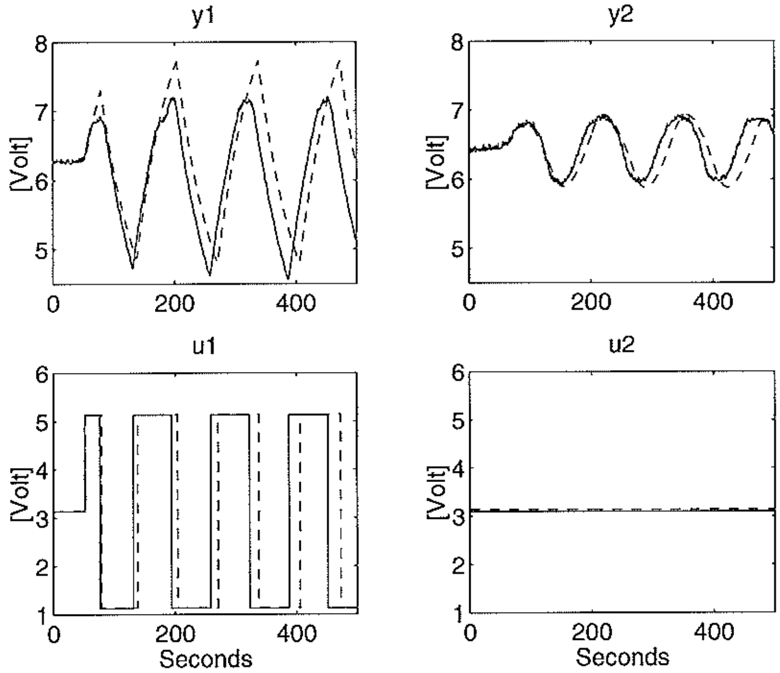

Figure 5 First loop under relay feedback. The simulated physical linear model(dashed) and the experiments (solid).

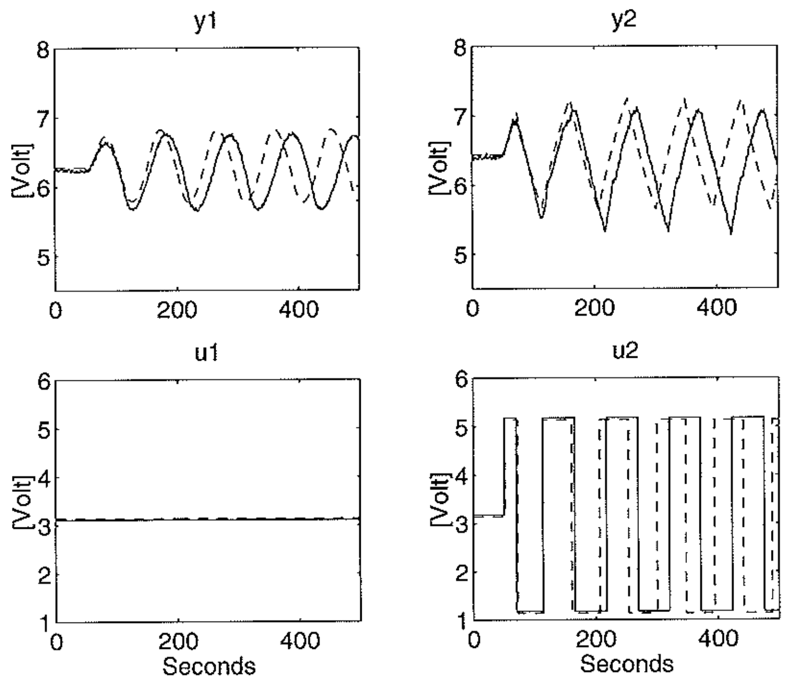

Figure 6 First loop under relay feedback. The simulated physical linear model(dashed) and the experiments (solid).

Summary As can be seen, the results from the simulation are very similar to the results from the experiments. The slightly difference can be explained by the fact that the simulated model is just a physical. Furthermore, the changing

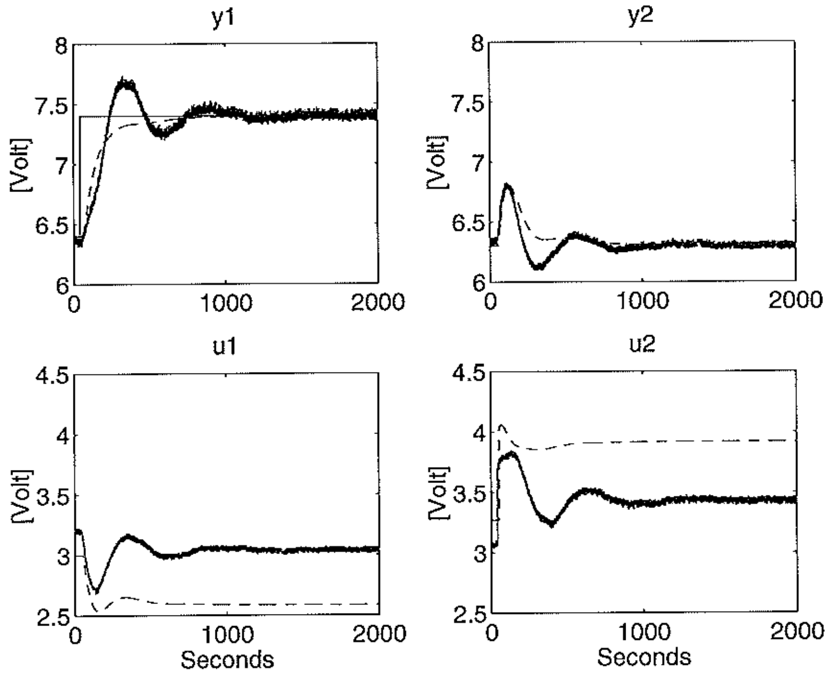

Figure 7 The results of the non minimum phase system controlled by two diagonal PI controllers. Same variables shown as in Figure 5.

of the configuration of the system, i.e., from the minimum phase system to non minimum phase system and vice versa, has been done a couple of times during the experiments. This was done by changing the position of the valves. Once we did that, it was very difficult to go back to the same old valve settings. All this implies that, the identified physical model may differ slightly from the real system.

It must be pointed out that, all the methods have been simulated and compared with the experiments and the results are of the same nature as the results shown above, meaning that there is a very good agreement between the simulations and experiments.

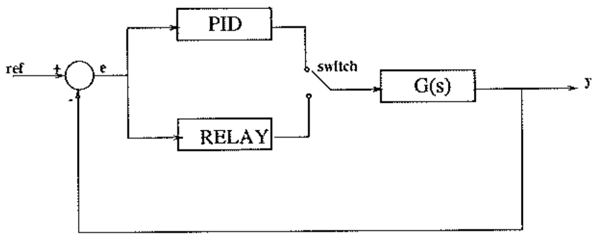

Figure 8 Auto-tuner of PID-controller.

3. Automatic Tuning of Controllers

3.1 Introduction

We start with a brief review of the relay-based auto-tuner for SISO systems. This method is based on a simple characterization of the system dynamics. The design is based on knowledge of the point on the Nyquist curve of the system transfer function $G(s)$ where the Nyquist curve intersects the negative real axis. This point is characterized by the parameters k_c and t_c, which are called the critical gain and the critical period. These parameters can be determined as in the original Ziegler-Nichols scheme [Ziegler and Nichols, 1943] or as in the method proposed by Åström and Hägglund [Åström and Hägglund, 1984].

Their method is based on the observation that a system with a phase lag of at least π at high frequencies may oscillate with period t_c under relay control. To determine the critical gain and the critical period, the system is connected in a feedback loop with a relay as is shown in Figure 8. The error e is then a periodic signal with the period t_c. If d is the relay amplitude, it follows from a Fourier series expansion that the first harmonic of the relay output has the amplitude $\frac{4d}{\pi}$. If the system output is a, the critical gain is thus approximately given by

$$k_c = \frac{4d}{\pi a}$$

This result also follows from the describing function $N(a)$

$$N(a) = \frac{4d}{\pi a}$$

A point on the Nyquist curve which is different from the critical point is obtained when the relay has hysteresis (ϵ). Then, the negative reciprocal of the describing function of such a relay is given by

$$\frac{-1}{N(a)} = -\frac{\pi}{4d}\sqrt{a^2 - \epsilon^2} - i\frac{\pi\epsilon}{4d}$$

This function is a straight line parallel to the real axis in the complex plane. See Figure 9. By choosing the relation between ϵ and d it is possible to determine a point on the Nyquist curve with a specified imaginary part. Several points on the Nyquist curve are easily obtained by repeating the experiment with different relations between ϵ and d.

A simple relay control experiment thus gives the information about the system which is needed in order to apply the design methods. This method

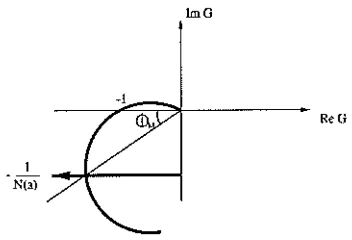

Figure 9 The negative reciprocal of the describing function and the Nyquist curve of $G(s)$.

has the advantage that it is easy to control the amplitude of the limit cycle by an appropriate choice of the relay amplitude. Another advantage is that the tuning experiment is executed under tight feedback control and that the experiment generates an input signal that is close to optimal for determining the critical point on the Nyquist curve.

To complete the description of the estimation method it is also necessary to give methods for automatic determination of the frequency and the amplitude of the oscillation.

The period of an oscillation can easily be determined by measuring the times between zero-crossings and the amplitude may be determined by measuring the peak-to-peak values.

There are some practical problems which must be solved before starting the tuning procedure, as for example we need to take care of the measurement noise because this can give errors in detection of peaks and zero-crossings. This can been done by introducing hysteresis in the relay.

When we had gone through the procedure above, it is straightforward to apply the classical Ziegler-Nichols tuning rules or some other tuning rules.

We will proceed and extend the SISO-relay feedback method to our multivariable system. In MIMO systems, the tuning problem is more complicated because of the interactions between loops. We will use a couple of various methods for relay tuning and compare the results.

3.2 SISO Controller Tuning

In the first method, relay experiment similar to the scalar experiment are applied: one loop is put under relay feedback while the second loop is kept open. Then, the second loop is put under relay feedback while the first loop is kept open. From the experiments above the controller parameters are then adjusted. See Figure 10.

Relay experiments

The tuning experiment is done in two phases. The first phase is an initial phase where we use a PID controller to move the system to the working point or we move it manually. When we reach the desired level we switch on the relay and begin our estimation procedure.

We have chosen the relay amplitude $d = 2.0[V]$ and relay hysteresis $\epsilon = 0.3[V]$.

Control design

Now, given the information about the critical point on the Nyquist curve $G(j\omega)$, it's possible to design the controller in such a way that we can move

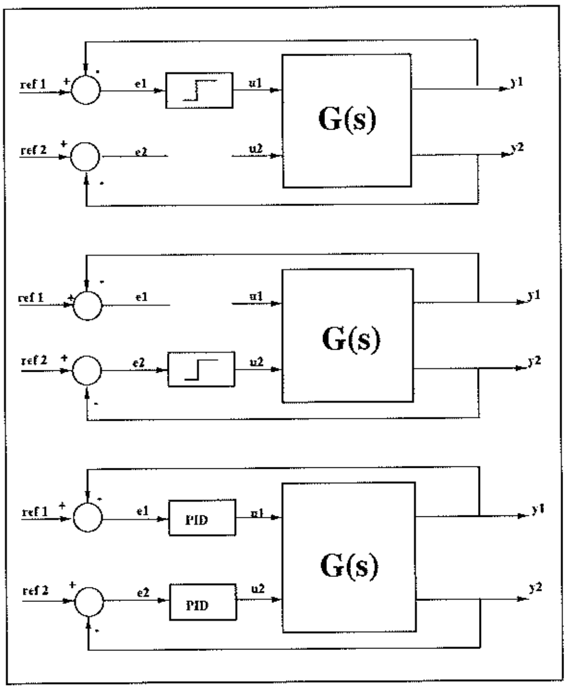

Figure 10 The algorithm for SISO controller tuning.

the Nyquist curve of the compensated system GC to a desired location at the frequency ω.

We will use a decentralized PID controller, which is one of the most common control schemes for interacting multiple-input multiple-output systems. The main reason for this is its relatively simple structure, which is easy to implement and to understand. The number of tuning parameters is $3n$, where n is the number of inputs and outputs, while in a full matrix there are $3n^2$ parameters. The decentralized controller C is block diagonal

$$C(s) = \begin{bmatrix} C_1(s) & 0 \\ 0 & C_2(s) \end{bmatrix}$$

with $C_1(s)$ and $C_2(s)$ parameterized as

$$K_{Pi}(1 + \frac{1}{T_{Ii}s} + T_{Di}s)$$

Our system, as we will see later in the performed experiments, behaves like a first order system. As a result of this we can solve the control problem with PI controllers. Derivative action is of little use for these problems. Furthermore, the noise will be amplified by the derivative term [Hägglund and Åström, 1991]. The parameters of the PI controllers are chosen as

$$K_{Pi} = 0.5a_1k_c$$

$$T_{Ii} = \frac{4a_2}{\omega}$$

where

$$k_c = \frac{4d}{\pi\sqrt{(a^2 - \epsilon^2)}}$$

and $a_1, a_2 \in [-1, 1]$ are constants.

The proposed approach is conceptually the same as Niederlinski's method [Niederlinski, 1971], which is based on the generalization of the classical ideas first presented by Ziegler and Nichols. This approach concentrates on getting a good design rather than the optimum one.

For the minimum phase system the parameters a_1 and a_2 are set equal to one. This means that, after computing the critical points, it is straightforward to compute the control parameters according to the suggestion given in [Hägglund and Åström, 1991].

For the non minimum phase system, control structure two, to be able to stabilize the system and achieve a reasonable settling time for a step response, we need to reduce the gain and have a larger integration time. After experimenting little with the constants a_1 and a_2, we arrive at the following values, $a_1 = 0.6$ and $a_2 = 2$, which give reasonable closed-loop performance.

When it comes to design the controller for the non minimum phase system, control structure one, then it's very hard to find good controller parameters. Moreover, because the $\det G(0) < 0$, it can be shown that the system is not DIC (Decentralized Integral Controllable), see Theorem 14.3-1 in [Morari and Zafirou, 1989]. This means that the closed loop system is not stable. Therefore, there exists no multi-loop PI controller with $K_{P1} = K_{P2} > 0$ that stabilizes the system.

In the same way as before, after little of experimenting with the constants a_1 and a_2, we arrive at the following values, $a_1 = 0.2$ and $a_2 = 7.5$ in the first loop and $a_1 = -0.015$ and $a_2 = 10$ in the second loop. The chosen values stabilize the system and gives reasonable settling times.

The period of the oscillation is determined by measuring the time between zero-crossings and the amplitude by measuring the peak-to-peak values of the output.

Validation

Minimum Phase System The results of relay experiments for the minimum phase system are shown in Figure 11 and Figure 12.

The system behaves like a first order system which can be seen from the relay experiment plots. The output of the system switches direction just after it passes the hysteresis. As a result of this, the experiment depends very much from the choice of the value of the hysteresis. This means that, depending on the value of relay hysteresis we choose, we will have different critical points which in turn lead to different settings, and hence different performance.

From the relay plots we can also see that the interaction is not so large which means that it shouldn't be so difficult to control this system by two diagonal PID controllers.

The controller parameters found for the minimum phase system are $(K_{P1}, T_{I1}) = (7.7, 13.4)$ and $(K_{P2}, T_{I2}) = (4.0, 12.7)$. They give quite good performance as shown in Figure 13, where the responses are given for a step change of the reference one. The system is well damped, with an overshoot of output y_1 approximately just below 12% and settling time of about 55 seconds.

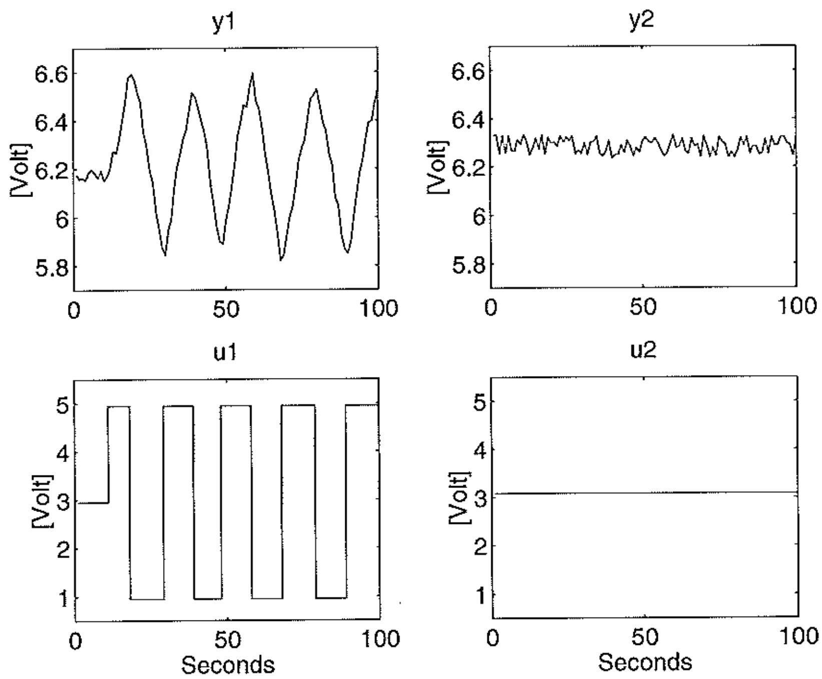

Figure 11 First loop under relay feedback (minimum phase system).

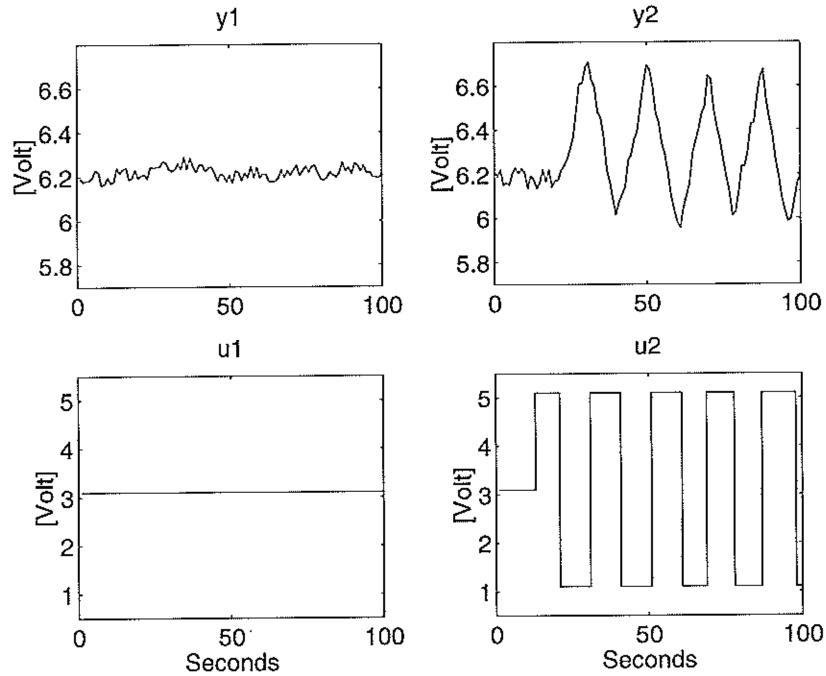

Figure 12 Second loop under relay feedback (minimum phase system).

We can also see that output y_2 remained almost unchanged when we made a step response for output y_1.

Non-Minimum Phase System, control structure one The results of relay experiments for the non minimum phase system paired as in the first

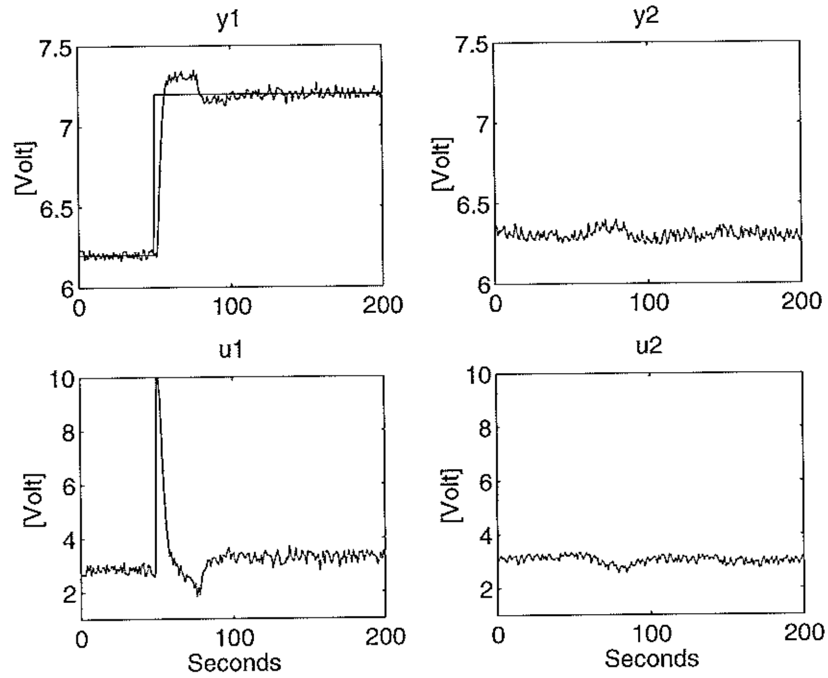

Figure 13 The results of the minimum phase system controlled by two diagonal PI controllers.

control structure (see Section 2.6) are shown in Figure 14 and Figure 15.

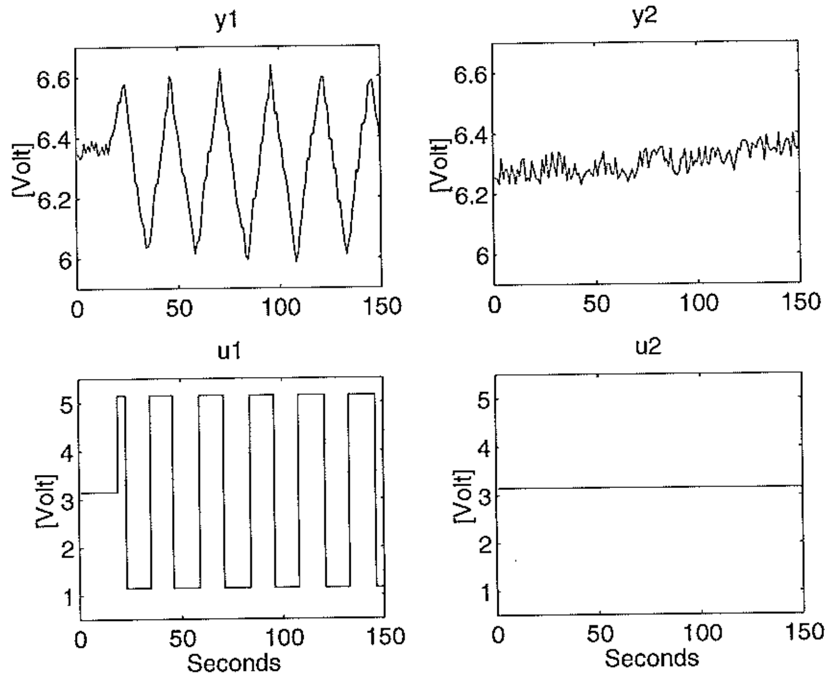

Figure 14 First loop under relay feedback (non minimum phase system, control structure one).

The controller parameters found for the non minimum phase system (control structure one) are $(K_{P1}, T_{I1}) = (1.4, 120.4)$ and $(K_{P2}, T_{I2}) = (-0.14, 195.6)$.

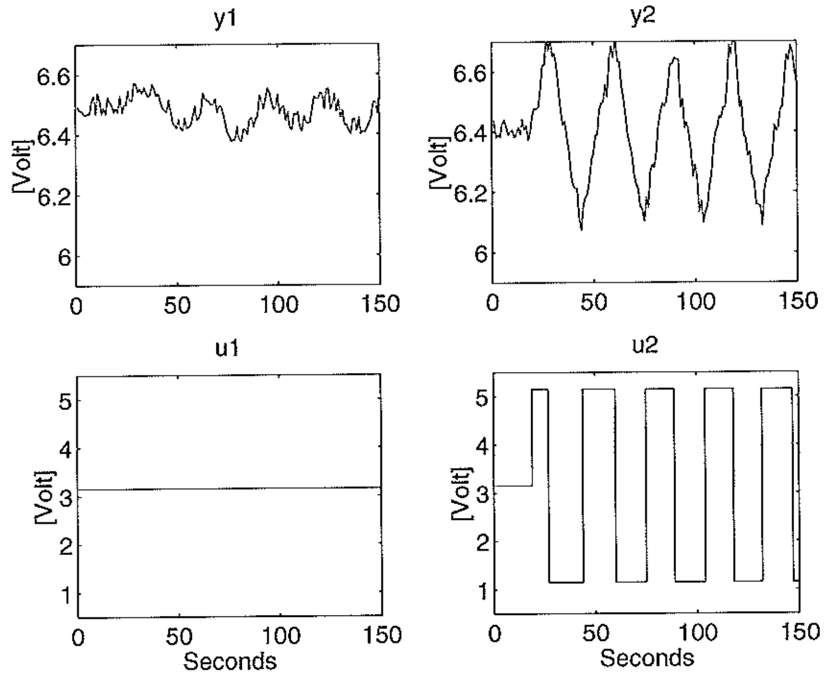

Figure 15 Second loop under relay feedback (non minimum phase system, control structure two).

The controller performance is shown in Figure 16, where the responses are given for a step change of the reference one. The system is stabilized, but it is very much slower compared to the minimum phase system. Notice that the settling time is approximately ten times longer. In addition to that a reference change in one loop affects the output in the other loop a great deal.

Non-Minimum Phase System, control structure two The results of relay experiments for the non minimum phase system paired as in the second control structure are shown in Figure 17 and Figure 18.

The controller parameters found are $(K_{P1}, T_{I1}) = (0.66\ 160.3)$ and $(K_{P2}, T_{I2}) = (0.75\ 145.2)$. The controller performance is shown in Figure 19, where the responses are given for a step change of the reference one. A reference change in one loop doesn't affect the output in the other loop in the same way as it did in control structure one.

3.3 Sequential Controller Tuning

The main idea of sequential relay tuning is to tune the multivariable system loop by loop, closing each loop once it is tuned, until all the loops are tuned [Vasnani, 1994]. To tune each loop, a relay feedback configuration is set up to determine the critical gain and frequency of the corresponding loop.

The sequence of loop closing is important since it affects the amount of interaction entering the first tuned loop and therefore limits the quality of the set point responses for those loops.

Relay experiments

Relay experiments similar to the method presented in Section 4.2 are applied: one loop is put under relay feedback and it's controller parameters are ad-

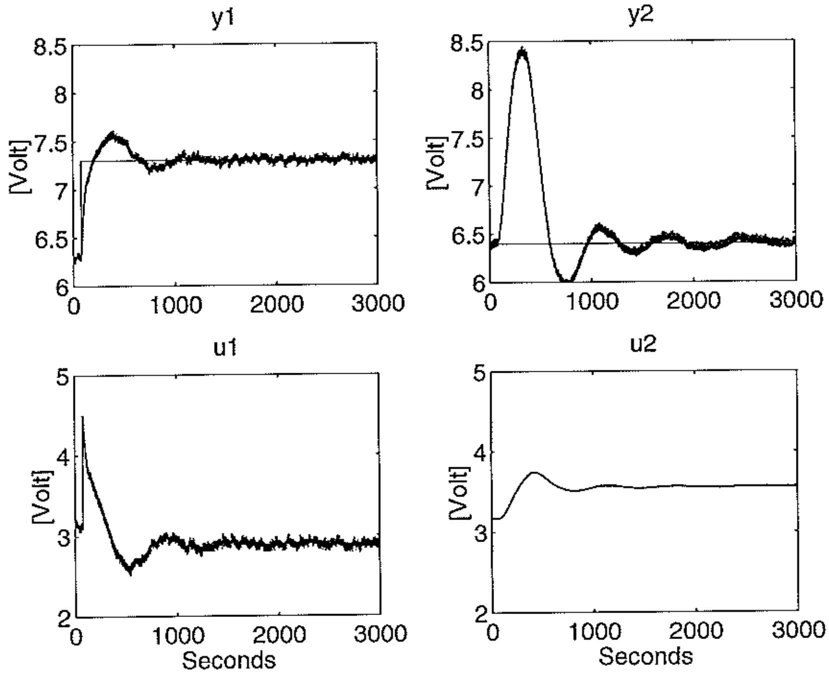

Figure 16 The results of the non minimum phase system (control structure one) controlled by two diagonal PI controllers.

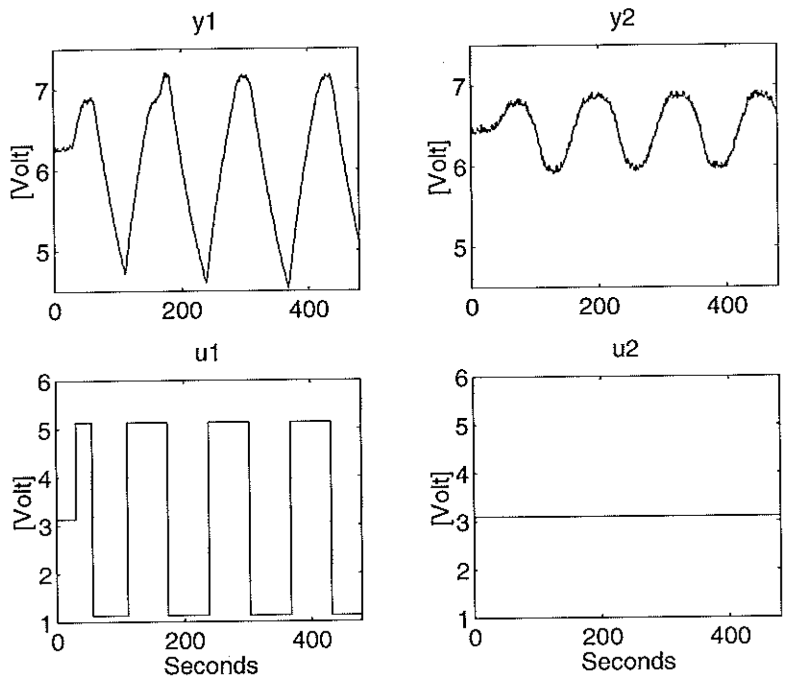

Figure 17 First loop under relay feedback (non minimum phase system, control structure two).

justed, then another loop is put under relay feedback while the first loop is controlled by a PID controller. The second loop parameters are then adjusted. See Figure 20.

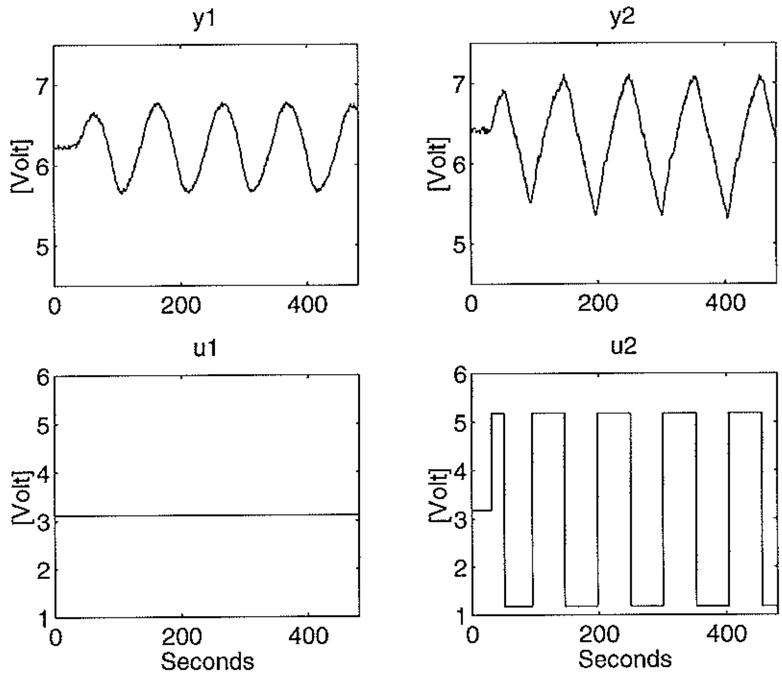

Figure 18 Second loop under relay feedback (non minimum phase system, control structure two).

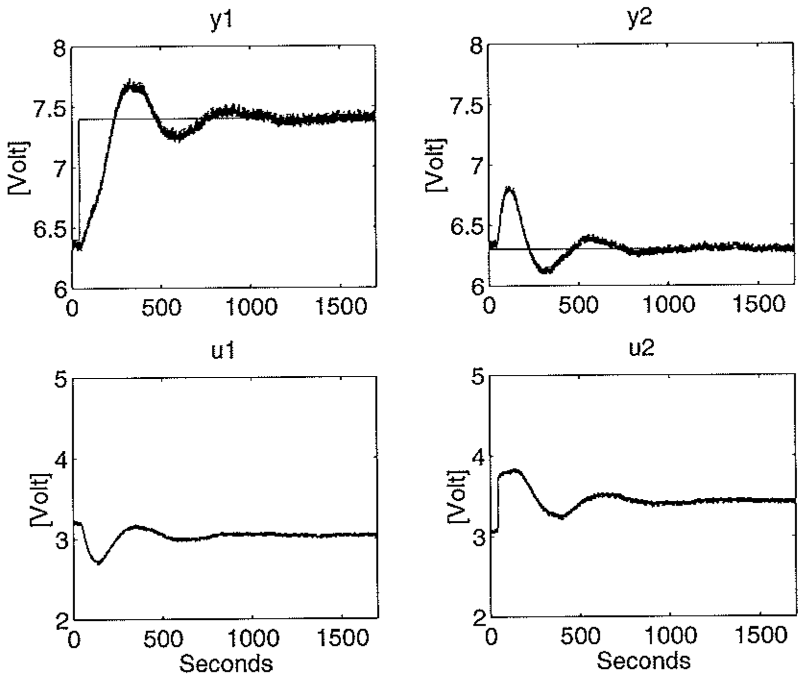

Figure 19 The results of the non minimum phase system (control structure two) controlled by two diagonal PI controllers.

The tuning experiment is done in two phases as before. The first phase is an initial phase where we use a PID controller to move the system to the working point or we do it manually. When we reach the desired level we switch

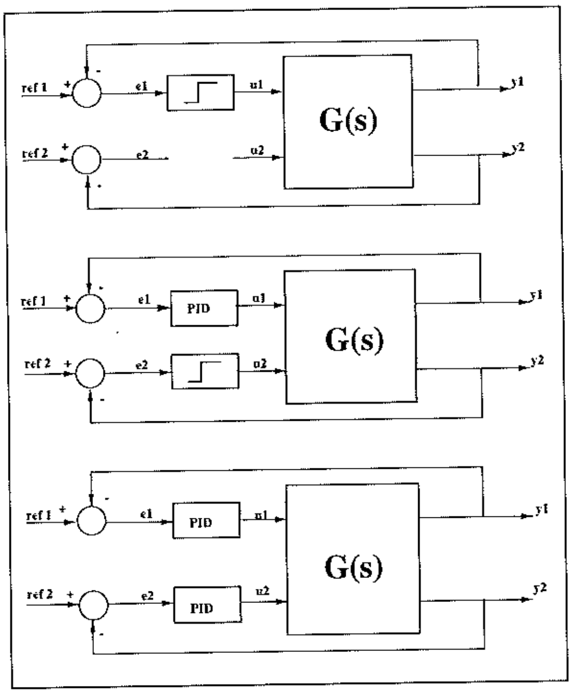

Figure 20 The algorithm for Sequential controller tuning.

on the relay and begin our estimation procedure.

The relay parameters have been chosen as in Section 4.2.

Control design

Given the information about the critical point on the Nyquist curve $G(j\omega)$, we design a decentralized controller, with the control parameters chosen in the same way as proposed in Section 4.2.

Validation

Minimum Phase System The result of relay experiment for the minimum phase system is shown in Figure 21 (only for the second loop, because the first loop experiment is the same as in the SISO Tuning).

Notice now that the input $u1$ is no longer constant. Instead the input $u1$ tries to make output $y1$ follow the reference signal, meaning that it minimizes the effect of the interaction. But, because the interaction is not so large, the benefits with the method will be negligible.

The controller parameters found for the minimum phase system are $(K_{P1}, T_{I1}) = (7.7, 13.4)$ and $(K_{P2}, T_{I2}) = (5.8, 14.6)$. They give slightly better result compared to the SISO method as shown in Figure 22. The system is well damped, with an overshoot of output y_1 approximately just below 10% and settling time of about 50 seconds.

Non-Minimum Phase System, control structure one The result of relay experiment for the non minimum phase system, control structure one, is shown in Figure 23.

The controller parameters found are $(K_{P1}, T_{I1}) = (1.4, 120.4)$ and $(K_{P2}, T_{I2}) = (-0.14, 226.6)$. They give slightly better result compared to the

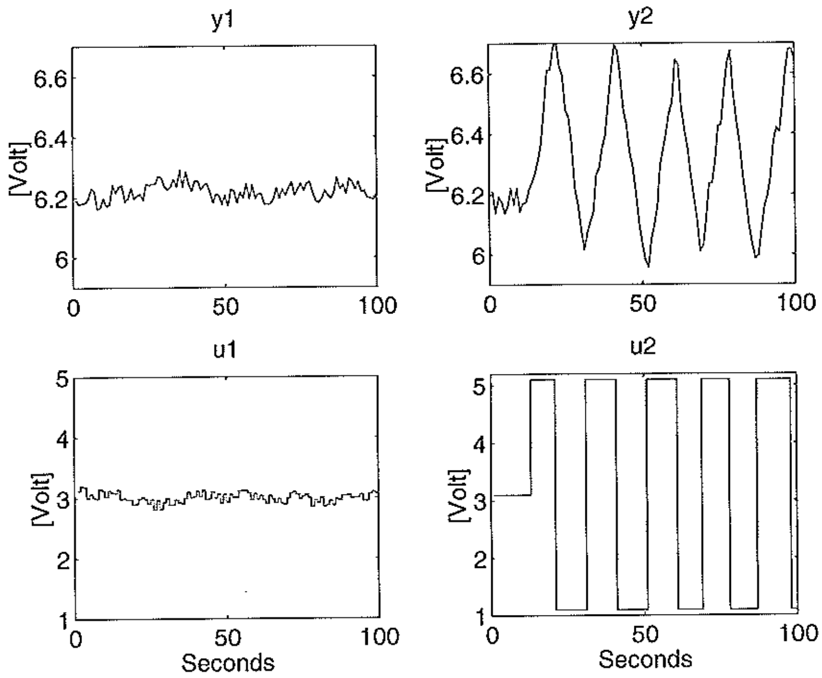

Figure 21 First loop controlled by a PI controller and the second loop under relay feedback (minimum phase system).

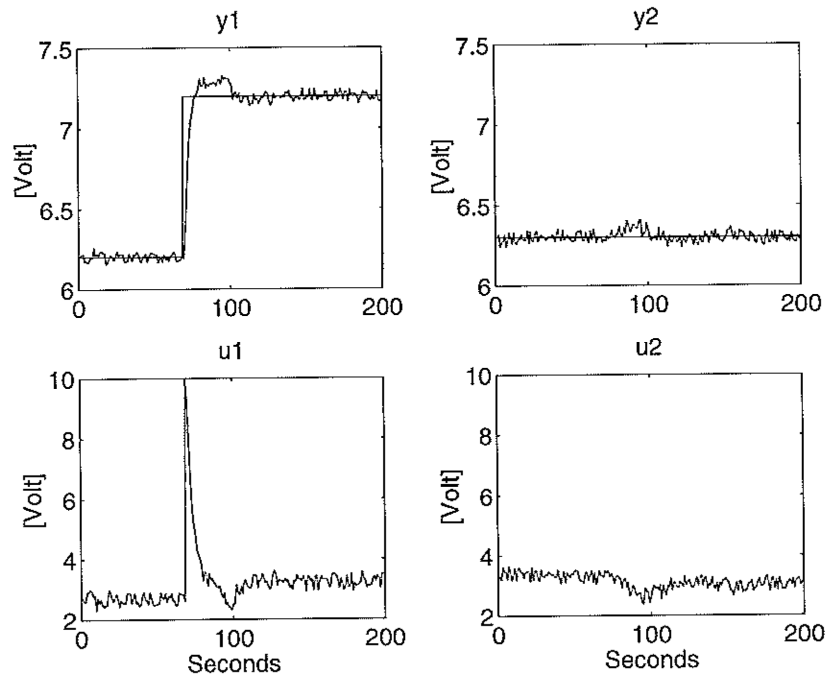

Figure 22 The results of the minimum phase system controlled by two diagonal PI controllers. The four plots show the experimental results.

SISO method as shown in Figure 24.

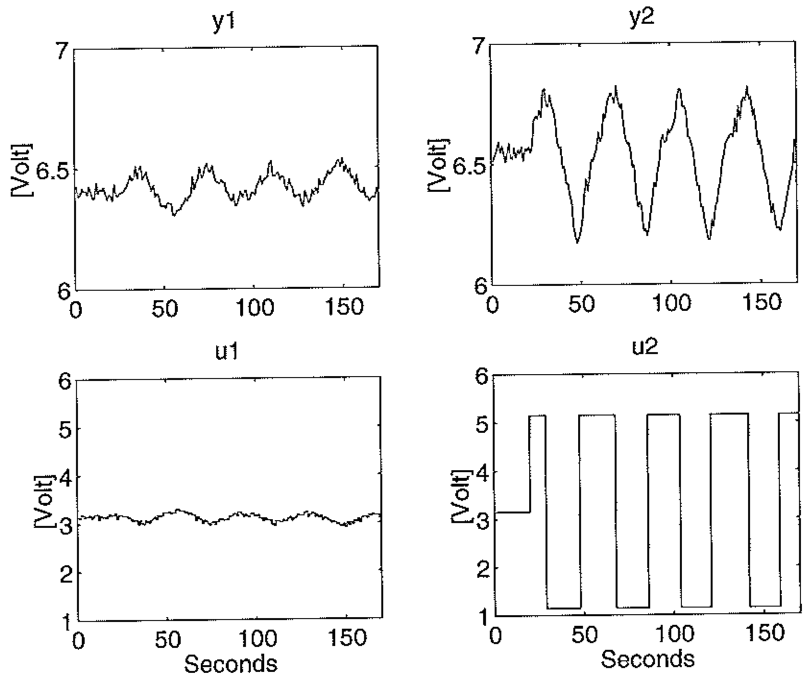

Figure 23 First loop controlled by a PI controller and the second loop under relay feedback (non minimum phase system, control structure one).

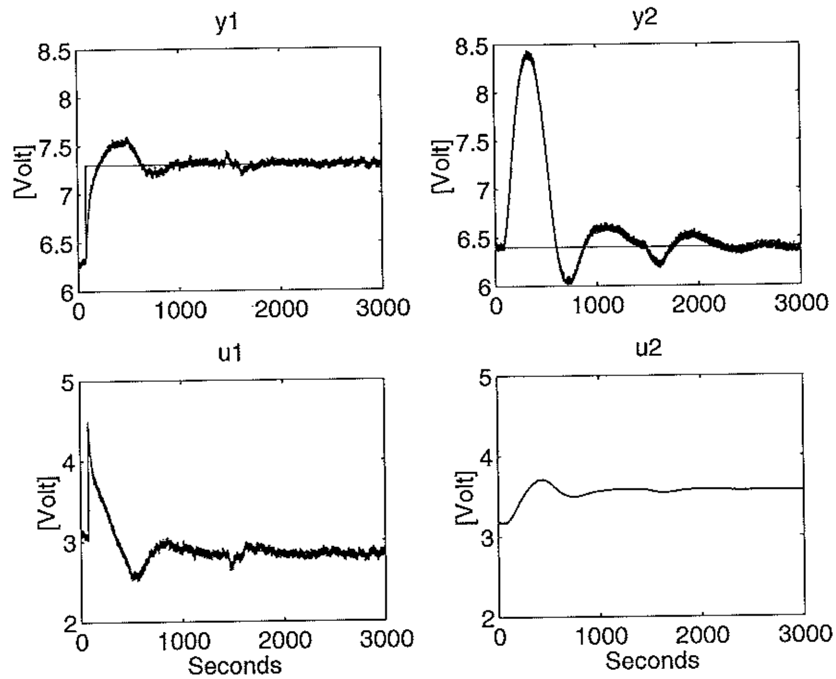

Figure 24 The results of the non minimum phase system (control structure one) controlled by two diagonal PI controllers.

Non-Minimum Phase System, control structure two The result of relay experiment for the non minimum phase system, control structure two, is shown in Figure 25.

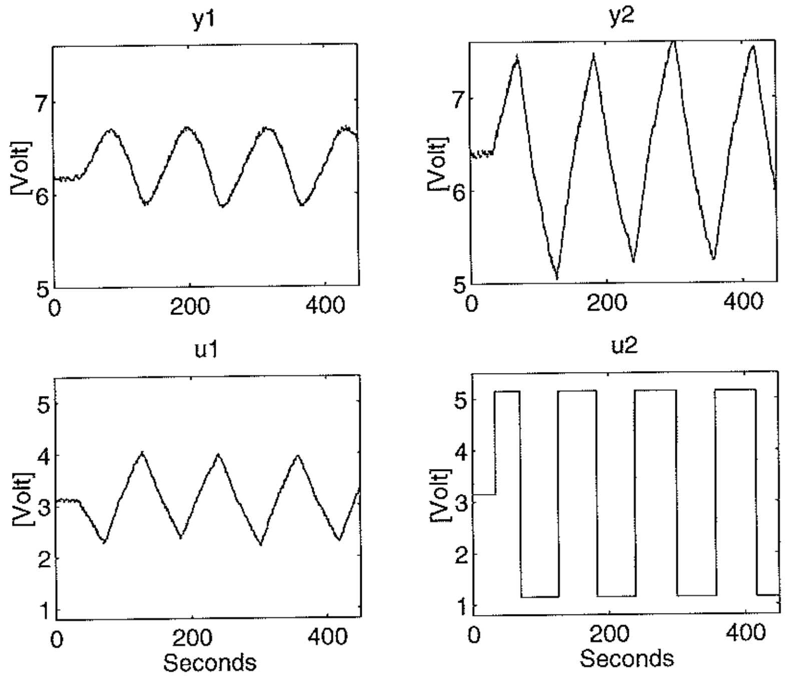

Figure 25 First loop controlled by a PI controller and the second loop under relay feedback (non minimum phase system, control structure two).

The controller parameters found for the non minimum phase system are $(K_{P1}, T_{I1}) = (0.66, 160.3)$ and $(K_{P2}, T_{I2}) = (0.61, 150.1)$. In this case too, the results are slightly better compared to the SISO methods, as shown in Figure 26.

3.4 Extension to Sequential Controller Tuning

A drawback with the original relay feedback experiment is its lack of excitation. Because only a square-wave of a single frequency enters the system, only models such as

$$G(s) = \frac{K}{1 + sT} e^{-sL}$$

can be estimated. If we are interested in more complex models we can do a modification of the standard relay experiment, by simply estimating three points on the Nyquist curve [Johansson, 1997].

The method is based on the observation that, a system with a relay in a series with a filter, will give any point on the Nyquist curve. Three crucial points are marked with crosses in Figure 27. Point 1 is determined by a standard relay experiment, point 0 is determined from a step-response experiment and point 2 is determined from an experiment with a relay and an integrator in series.

The aim of this tuning method is to improve the performance of one loop by adjusting appropriate elements of the controller matrix. For doing that we need at least the knowledge of the SIMO transfer matrix from the input signal of the badly tuned loop to the outputs of the system.

The system outputs can be described by the following equations:

$$y_1 = G_{11} u_1 + G_{12} u_2$$

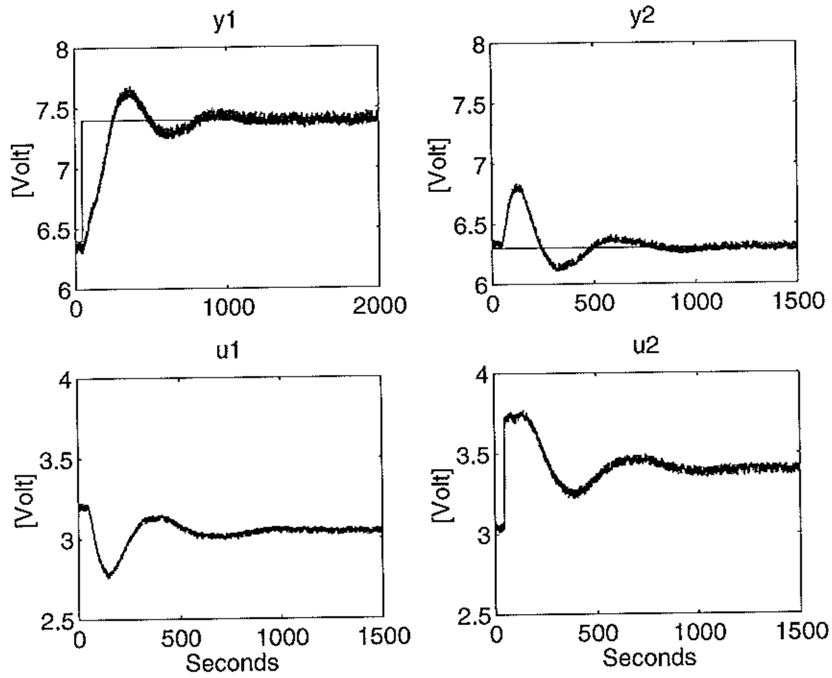

Figure 26 The results of the non minimum phase system (control structure two) controlled by two diagonal PI controllers.

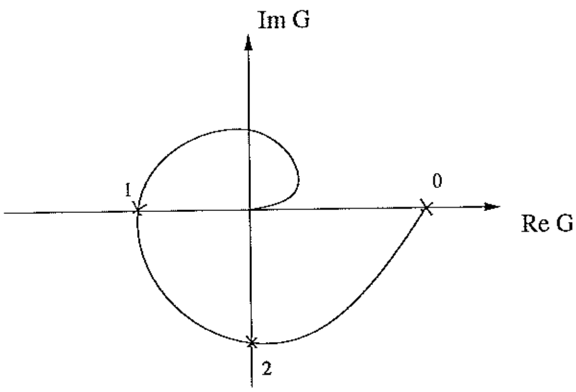

Figure 27 Three important points on the Nyquist curve.

$$y_2 = G_{21}u_1 + G_{22}u_2$$

where

$$u_1 = -C_1 y_1$$

$$u_2 = -C_2 y_2$$

If we insert u_1 in y_1 we will have

$$y_1 = \frac{G_{12}}{1 + G_{11}C_1} u_2$$

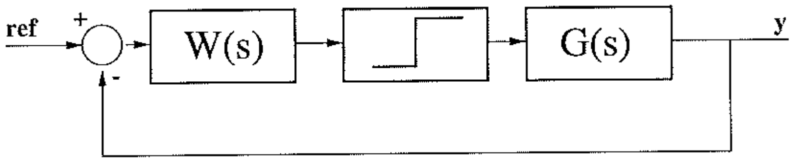

Figure 28 Extended relay experiment.

Let H_1 then be defined as

$$H_1 = \frac{G_{12}}{1 + G_{11}C_1}$$

In the same way as before, we insert u_1 and the derived equation for y_1 in y_2, and after some calculations we have

$$y_2 = (\frac{G_{12}G_{21}C_1}{1 + G_{11}C_1} - G_{22})u_2$$

Let H_2 then be defined as

$$H_2 = \frac{G_{12}G_{21}C_1}{1 + G_{11}C_1} - G_{22}$$

Relay experiments

Figure 28 shows the extended relay experiment applied to the MIMO system.

The algorithm for identifying the three points is as follows:

1. Set $W = 1$ and wait for a stationary oscillation. Measure the frequency ω_1 and derive the response of H_1 and H_2.
2. Set $W = \frac{1}{s}$ and wait for a stationary oscillation. Measure the frequency ω_2 and derive the response of H_1 and H_2.
3. Freeze the relay output and wait for steady-state and derive the steady-state gains for H_1 and H_2.
4. Estimate H as

$$G(s) = \frac{b_0 s + b_1}{s^3 + a_1 s^2 + a_2 s + a_3}$$

based on the responses and the corresponding frequencies ω_1 and ω_2.

Control design

For our quad tank system, as we will see later from the results, despite that we have estimated a more complex model, we can't expect to design a better controller than in the previous methods. The reason is that the two identified points (point 1 and point 2) are very close to each other. The identified model in this way is not so much better than the model identified before, and thus we can design the controller in the same way as we did in the previous methods.

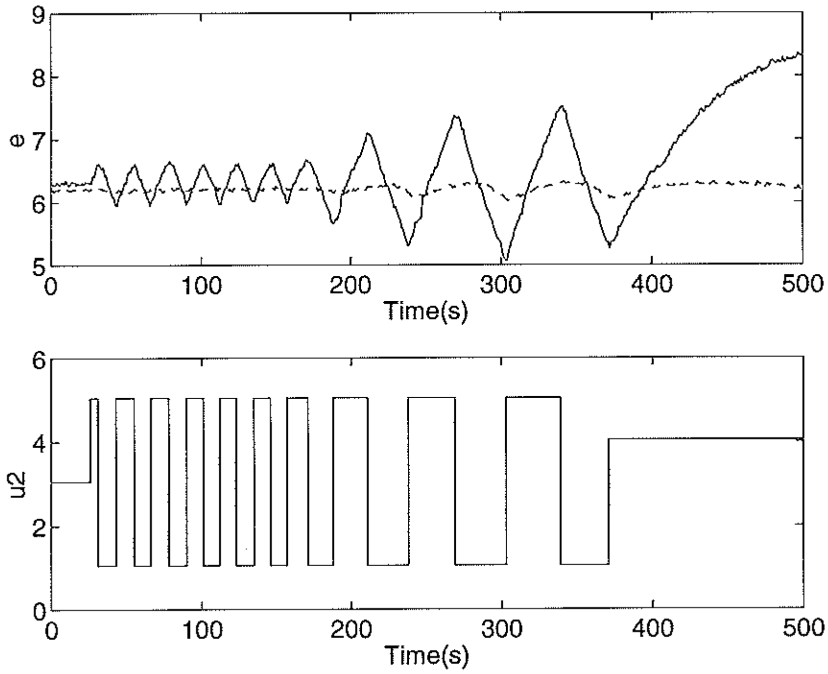

Figure 29 Extended relay experiment for the minimum phase system. The error signal e_1(dashed) is negligible compared to e_2(solid).

Validation

Minimum Phase System Figure 29 shows the response of the extended relay experiment for the minimum phase system. As can be seen, the response of e_1 is small compared to e_2.

The identified frequency points for H_2 from the experiment are shown in Figure 30. The result coincides very good with the expected ones. We also compute manually the phase difference so that we may double check the results. The phase difference for point 1 is approximately -3.0518 and the phase difference for point 2 is approximately -1.9539, which seem to be quite reasonable. This information we use then to estimate a third-order model of H_2.

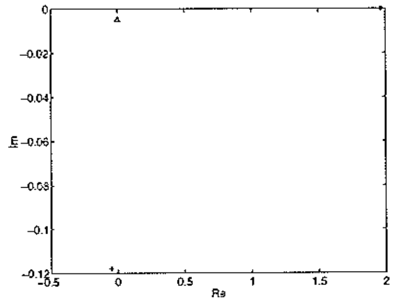

Figure 30 Frequency points for H_2. Point 1 is marked $+$, point 2 is marked $\wedge$ and point 0 is marked $*$.

The result of this estimation is presented in Figure 31, where we have drawn the Nyquist curve of the estimate of H_2.

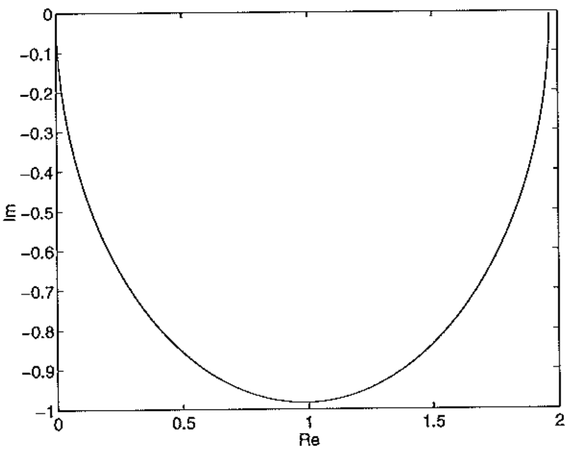

Figure 31 Nyquist curve of estimated H_2 for the minimum phase system.

The result from the extended relay experiment indicates that we can neglect the influence of H_1 and simply re-tune the controller K_2.

Non-Minimum Phase System The same procedure as above is done for the non minimum phase system, control structure one and control structure two.

Figure 32 shows the response of the extended relay experiment.

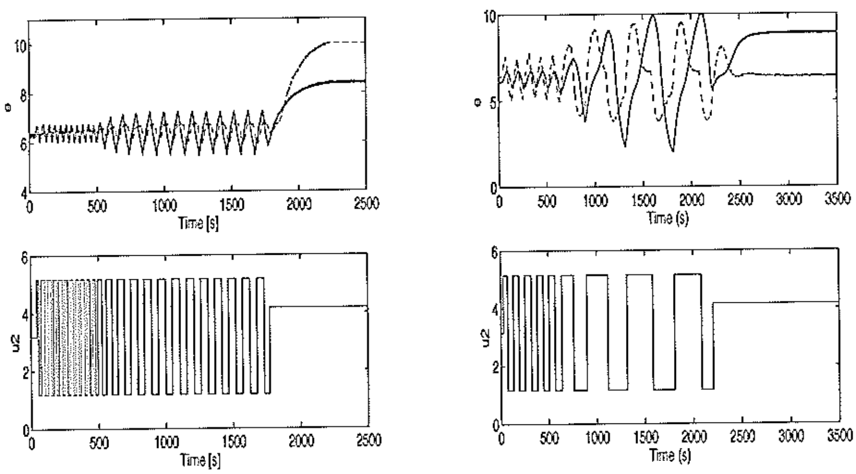

Figure 32 Extended relay experiment for the non minimum phase system (control structure one on the left and control structure two on the right). The error signal e_1(dashed) and e_2(solid).

This information we use then to estimate a third-order model of H_2 and H_1. The result of this estimation is presented in Figure 33, were we have drawn the Nyquist curve of the estimate of H_2(solid) and H_1(dashed).

Note that, for control structure two, the estimated point 0 from the experiment is insignificant. Hence the Nyquist diagram of estimated curve H_1 is negligible compared to the estimated curve H_2, and that is the reason for not seeing the H_1 curve.

The result from the extended relay experiment indicates that the interaction is significant compared with the minimum phase system, so it's probably not enough to re tune only the second loop K_2.

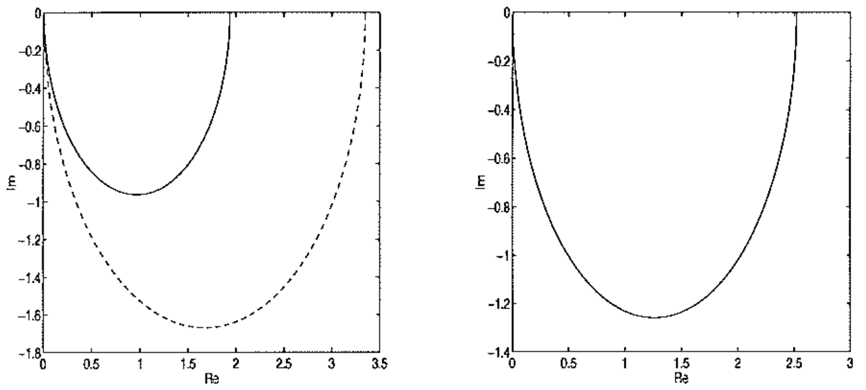

Figure 33 Nyquist curve of estimated H_2(solid) and H_1(dashed) for the non minimum phase system (control structure one on the left and control structure two on the right).

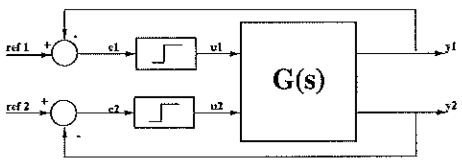

Figure 34 Decentralized relay test.

3.5 Decentralized Controller Tuning

This is a method for automatically tuning fully cross-coupled multivariable PID controllers from decentralized relay feedback [Wang et al., 1997].

Auto-tuning of controllers requires knowledge of the desired critical point, which consists of the critical gain and the critical frequency. For this purpose introduce a nonlinear feedback of the relay type in order to generate a limit cycle oscillation. Relay feedback is a simple and reliable test that keeps the system output under closed-loop control and makes it close to the operating point. This tuning device has one parameter that must be specified in advance, namely, the initial amplitude of the relay. It is also useful to introduce hysteresis in the relay. This reduces the effects of measurement noise and also increases the period of the oscillation. With hysteresis there is one more parameter which can be set based on a determination of the measurement noise level. A bias is also introduced into the relay to additionally obtain the system steady-state matrix. Then the controller is designed, where a new set of design equations are derived.

Relay experiments

It is found that for typical coupled multivariable systems, m outputs normally have the same oscillation frequencies. We want here to estimate the system frequency response $G(j\omega)$ at the critical oscillation frequency ω_c for controller tuning. To identify the steady-state gain matrix of the system additionally, a biased relay is used in the dominant loop to make the system inputs and outputs have non-zero means. Then one waits for the system to reach stationarity. The system stationary inputs $u_1(t)$, $u_2(t)$ and outputs $y_1(t)$, $y_2(t)$ are all periodic and they can be expanded in Fourier series. If the oscillations in two loops have a common frequency ω_c then the direct-current components and the first harmonics of these periodic waves are extracted as

$$U^1(0) = \begin{bmatrix} \int_0^{T_c} u_1(t)dt \\ \int_0^{T_c} u_2(t)dt \end{bmatrix}$$
$$Y^1(0) = \begin{bmatrix} \int_0^{T_c} y_1(t)dt \\ \int_0^{T_c} y_2(t)dt \end{bmatrix}$$

$$U^1(j\omega_c) = \begin{bmatrix} \int_0^{T_c} u_1(t)e^{-j\omega_c t}dt \\ \int_0^{T_c} u_2(t)e^{-j\omega_c t}dt \end{bmatrix}$$
$$Y^1(j\omega_c) = \begin{bmatrix} \int_0^{T_c} y_1(t)e^{-j\omega_c t}dt \\ \int_0^{T_c} y_2(t)e^{-j\omega_c t}dt \end{bmatrix}$$

Then we have

$$Y^1(0) = G(0)U^1(0)$$
$$Y^1(j\omega_c) = G(j\omega_c)U^1(j\omega_c)$$

Since the above equations are vector equations, they are not sufficient to determine $G(j\omega_c)$ and $G(0)$ from Y^1 and U^1 only. So, we need to do another experiment where we increase the relay amplitude of the dominant loop or decrease that of the another loop.

$$U^2(0) = \begin{bmatrix} \int_0^{T_c} u_1(t)dt \\ \int_0^{T_c} u_2(t)dt \end{bmatrix}$$
$$Y^2(0) = \begin{bmatrix} \int_0^{T_c} y_1(t)dt \\ \int_0^{T_c} y_2(t)dt \end{bmatrix}$$

$$U^2(j\omega_c) = \begin{bmatrix} \int_0^{T_c} u_1(t)e^{-j\omega_c t}dt \\ \int_0^{T_c} u_2(t)e^{-j\omega_c t}dt \end{bmatrix}$$
$$Y^2(j\omega_c) = \begin{bmatrix} \int_0^{T_c} y_1(t)e^{-j\omega_c t}dt \\ \int_0^{T_c} y_2(t)e^{-j\omega_c t}dt \end{bmatrix}$$

We have obtained

$$[Y^1(0)Y^2(0)] = G(0)[U^1(0)U^2(0)]$$
$$[Y^1(j\omega_c)Y^2(j\omega_c)] = G(j\omega_c)[U^1(j\omega_c)U^2(j\omega_c)]$$

From the above equations follows that the steady-state gain matrix G(0) and frequency-response matrix $G(j\omega_c)$ are determined as

$$G(0) = [Y^1(0)Y^2(0)][U^1(0)U^2(0)]^{-1}$$
$$G(j\omega_c) = [Y^1(j\omega_c)Y^2(j\omega_c)][U^1(j\omega_c)U^2(j\omega_c)]^{-1}$$

From the above we can say that our tuning experiment consist of two decentralized relay tests. For the first test, the relay amplitude for each loop is

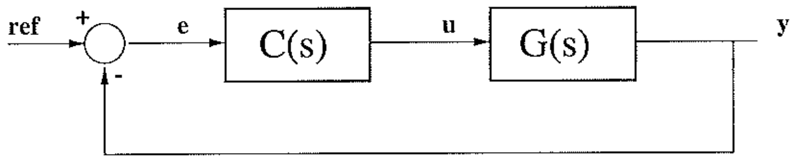

Figure 35 Multivariable control system.

set as in the single-variable case. For the second test, either the amplitude in the dominant loop is increased or the amplitude in other loop is decreased by 5-20%.

It is important that the system outputs oscillates with the same frequency. Furthermore, the system outputs must be asymmetric, because otherwise $U^1(0)$, $U^2(0)$, $Y^1(0)$ and $Y^2(0)$ are all equal to zero.

Control design

Having the frequency response available at two points, namely, $\omega = 0$ and ω_c, we can consider a multivariable control system as shown in Figure 35. The objectives of control design are to make the closed-loop control system decoupled and the resultant independent loops have good transient and accuracy in the usual sense of single variable systems. For the system in Figure 35 to be decoupled, it is shown by [Wang, 1992] that the open-loop transfer matrix $Q = GC$ must be diagonal and nonsingular. For each column of GC we have

$$G_{11}C_{11} + G_{12}C_{21} = Q_{11} \quad (7)$$
$$G_{21}C_{12} + G_{22}C_{22} = Q_{22} \quad (8)$$

$$G_{21}C_{11} + G_{22}C_{21} = 0$$
$$G_{11}C_{12} + G_{12}C_{22} = 0$$

Introduce f_{21} and f_{12}

$$f_{21} = -\frac{G_{21}}{G_{22}}$$
$$f_{12} = -\frac{G_{12}}{G_{11}}$$

then we can obtain C_{21} and C_{12} in terms of C_{11} and C_{22} as

$$C_{21} = f_{21}C_{11}$$
$$C_{12} = f_{12}C_{22}$$

Equation 7 and equation 8 then becomes

$$\tilde{G}_{11}C_{11} = Q_{11} \quad (9)$$
$$\tilde{G}_{22}C_{22} = Q_{22} \quad (10)$$

where

$$\tilde{G}_{11} = G_{11} + G_{12}f_{21}$$
$$\tilde{G}_{22} = G_{22} + G_{21}f_{12}$$

One notes that equation 9 and equation 10 are independent of the controller off-diagonal elements, but contains only the controller diagonal elements C_{11} and C_{22}. This can be designed for the equivalent single-variable plant $\tilde{G}_{11}$ and $\tilde{G}_{22}$ with a single-variable design method.

Given the information on the system as $\tilde{G}_{11}(0)$, $\tilde{G}_{22}(0)$, $\tilde{G}_{11}(j\omega_c)$ and $\tilde{G}_{22}(j\omega_c)$ there are many SISO methods available to tune k_{11} and k_{22} of PID type. It turns out that the gain and phase margin rule is most suitable for our application, since performance and robustness of the system thus tuned is very promising for most cases.

With the method in [Ho et. al, 1995], the points $\tilde{G}_{11}(0)$, $\tilde{G}_{22}(0)$ and $\tilde{G}_{11}(j\omega_c)$, $\tilde{G}_{22}(j\omega_c)$ are fitted to a first-order plus dead-time model

$$\tilde{G}_{ii}(s) = [\frac{\tilde{C}_{ii}}{1 + s\tilde{\tau}_{ii}}]e^{-s\tilde{L}_{ii}}$$

Let us derive the equations for $\tilde{L}_{ii}$, $\tilde{C}_{ii}$ and $\tilde{\tau}_{ii}$.

Let $s = 0$, then

$$\tilde{G}_{ii}(0) = \tilde{C}_{ii}$$

Let $s = j\omega_c$, then

$$\tilde{G}_{ii}(j\omega_c) = [\frac{\tilde{C}_{ii}}{1 + j\omega_c\tilde{\tau}_{ii}}]e^{-j\omega_c\tilde{L}_{ii}}$$

$$\mid \tilde{G}_{ii}(j\omega_c) \mid = [\frac{\tilde{C}_{ii}}{1 + j\omega_c\tilde{\tau}_{ii}}]$$

and we get

$$\tilde{\tau}_{ii} = \frac{1}{\omega_c}\sqrt{\frac{\tilde{C}_{ii}^2}{\mid \tilde{G}_{ii}(j\omega_c) \mid^2} - 1}$$

In the same way we derive the following result

$$\tilde{L}_{ii} = \frac{1}{\omega_c}(-arg[\tilde{G}_{ii}(j\omega_c)] - tan^{-1}(\omega_c\tilde{\tau}_{ii}))$$

The controllers diagonal elements C_{ii} are taken as

$$C_{ii}(s) = K_{Pii}(1 + 1/T_{Iii}(s))$$

and its parameters are given by

$$K_{Pii} = \frac{\omega_{Pii}\tilde{\tau}_{ii}}{A_M\tilde{C}_{ii}}$$

$$T_{Iii} = (2\omega_{Pii} - \frac{4\omega_{Pii}^2\tilde{L}_{ii}}{\pi} + \frac{1}{\tilde{\tau}_{ii}})^{-1}$$

where

$$\omega_{Pii} = \frac{A_M\Phi_M + 0.5\pi A_M(A_M - 1)}{(A_M^2 - 1)\tilde{L}_{ii}}$$

and A_M, Φ_M are the specified gain and phase margins respectively.

For the off-diagonal elements we choose C_{ij} of PID type as

$$C_{ij}(s) = K_{Pij}(1 + 1/T_{Iij}(s) + T_{Dij}(s))$$

such that the following is satisfied at $\omega = 0$ and $\omega = \omega_c$, i.e.,

$$\lim_{s\to 0} sC_{ij}(s) = \lim_{s\to 0} sf_{ij}(s)C_{jj}(s)$$

$$C_{ij}(j\omega_c) = f_{ij}(j\omega_c)C_{jj}(j\omega_c) =: \gamma_{ij}e^{j\varphi_{ij}}$$

We give a derivation of the controller parameters

$$\lim_{s\to 0} sC_{ij}(s) = \lim_{s\to 0} sf_{ij}(s)C_{jj}(s) = f_{ij}(0)\frac{K_{Pjj}}{T_{Ijj}}$$

$$\lim_{s\to 0} C_{ij}(s) = \lim_{s\to 0} sK_{Pij}(1 + 1/T_{Iij}(s) + T_{Dij}(s)) \to \frac{K_{Pij}}{T_{Iij}}$$

From the above equations we have that

$$T_{Iij} = \frac{K_{Pij}T_{Ijj}}{K_{Pjj}f_{ij}(0)}$$

Next we let $s = j\omega_c$. Then

$$C_{ij}(j\omega_c) = f_{ij}(j\omega_c)C_{jj}(j\omega_c) =: \gamma_{ij}e^{j\varphi_{ij}} = K_{Pij}(1 + 1/T_{Iij}(s) + T_{Dij}(s))$$

$$\gamma_{ij}(\cos\varphi_{ij} + j\sin\varphi_{ij}) = K_{Pij}(1 + 1/T_{Iij}(s) + T_{Dij}(s))$$

Identification of the left hand side with the right hand side gives us the following

$$K_{Pij} = \gamma_{ij}\cos\varphi_{ij}$$

$$T_{Dij} = \frac{1}{\omega_c}(tan\varphi_{ij} + \frac{1}{T_{Iij}\omega_c})$$

Validation

Now, we try to identify the system as proposed in the theory, but with a significant change in the identification procedure.

For the identification of the stationary values it was proposed in the theory that we should have a biased relay in the second loop. This is not possible for our quad tank system because the system behaves as a first order system. In general, for identifying the steady-state matrix we introduce a bias in the relay outputs and then we expect that the process' outputs are also of non-zero mean. This is not the case with our system and thus the method fail to identify the steady-state matrix.Therefore, we identify the stationary values of the system using another method, see item 4. This means that we still have to do two experiments for the identification as proposed in the theory, but the identification takes longer time to complete.

Minimum Phase System The algorithm for identifying the steady-state gain matrix $\hat{G}(0)$ and the frequency response matrix $\hat{G}(j\omega_c)$ is as follows.

The first test:

1. The relay in loop 1 is set as an ideal symmetric relay with switching levels 0.2 and -0.2.
2. The relay in loop 2 is set in the same way as in loop 1.
3. Wait for a stationary oscillation.
4. Freeze the relay output in loop 1 and wait for steady-state and derive the steady-state gains for $\hat{G}_{11}$ and $\hat{G}_{21}$.

The second test:

1. The relay in loop 1 is set as before with switching levels 0.2 and -0.2.
2. The switching levels of the relay in loop 2 are changed to 0.205 and -0.205.
3. Wait for a stationary oscillation.
4. Freeze the relay output in loop 2 and wait for steady-state and derive the steady-state gains for $\hat{G}_{22}$ and $\hat{G}_{12}$.

In the first test the system exhibits limit-cycle oscillations of common frequency with frequency $\omega_c^1 = 0.366$.

In the second test it exhibits a limit-cycle oscillations of common frequency, with $\omega_c^2 = 0.375$.

The steady-state gain matrix $\hat{G}(0)$ and the frequency-response matrix $\hat{G}(j\omega_c)$ are then computed. The results are

$$\hat{G}(0) = \begin{bmatrix} 2.66 & 1.46 \\ 1.37 & 2.87 \end{bmatrix}$$

$$\hat{G}(j\omega_c) = \begin{bmatrix} 0.1075e^{-1.79j} & 0.0040e^{-2.82j} \\ 0.0107e^{-3.13j} & 0.0830e^{-1.74j} \end{bmatrix}$$

where $\omega_c = 0.5(\omega_c^1 + \omega_c^2) = 0.37$. They are quite accurate compared with their "true" values:

$$G(0) = \begin{bmatrix} 2.6 & 1.5 \\ 1.4 & 2.8 \end{bmatrix}$$

$$G(j\omega_c) = \begin{bmatrix} 0.11e^{-1.52j} & 0.0076e^{-2.98j} \\ 0.0038e^{-3.02j} & 0.084e^{-1.54j} \end{bmatrix}$$

The interaction in the minimum-phase system is not so large, thus the outputs don't have the exact same oscillations frequencies. Our integral computations are very dependent on the fact that we have the same oscillations frequencies on the system outputs, thus the results differ slightly from the "true" ones.

Now we use the information that we have from the identification to design the controller derived before.

The gain margin and phase margin are chosen as $A_{m1} = A_{m2} = 3$ and $\phi_{m1} = \phi_{m2} = \frac{\pi}{3}$. Then yield the following controller

$$C(s) = \begin{bmatrix} 0.73(1+\frac{1}{46.2s}) & -0.005(1+\frac{1}{0.56s}-23s) \\ -9.3296e^{-4}(1+\frac{1}{0.11s}-30s) & 0.98(1+\frac{1}{64.5s}) \end{bmatrix}$$

The tuning processes and the resulting performance are shown in Figure 36, Figure 37 and Figure 38.

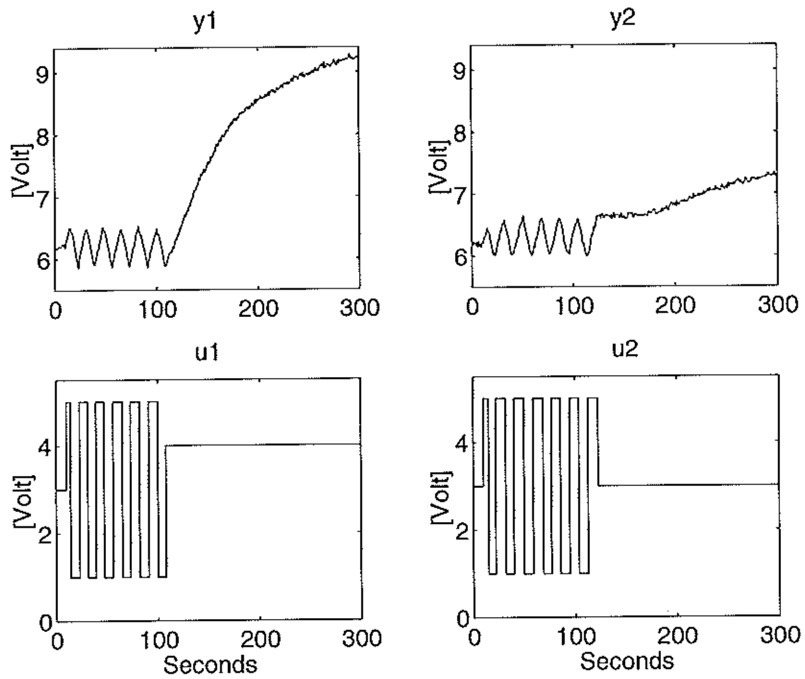

Figure 36 The first test applied to the minimum phase system.

Non-Minimum Phase System, control structure one We perform the same procedure as above for the non minimum phase system, control structure one.

The relay switching levels in both tests have been chosen as in the tests performed for the minimum phase system.

In the first test the system exhibits limit-cycle oscillations of common frequency with frequency $\omega_c^1 = 0.207$.

In the second test it exhibits a limit-cycle oscillations of common frequency, with $\omega_c^2 = 0.209$.

The steady-state gain matrix $\hat{G}(0)$ and the frequency-response matrix $\hat{G}(j\omega_c)$ for the this system are then computed. The results are

$$\hat{G}(0) = \begin{bmatrix} 1.56 & 3.21 \\ 2.93 & 1.62 \end{bmatrix}$$

$$\hat{G}(j\omega_c) = \begin{bmatrix} 0.107e^{-1.69j} & 0.0097e^{-3.21j} \\ 0.0085e^{-3.35j} & 0.096e^{-1.71j} \end{bmatrix}$$

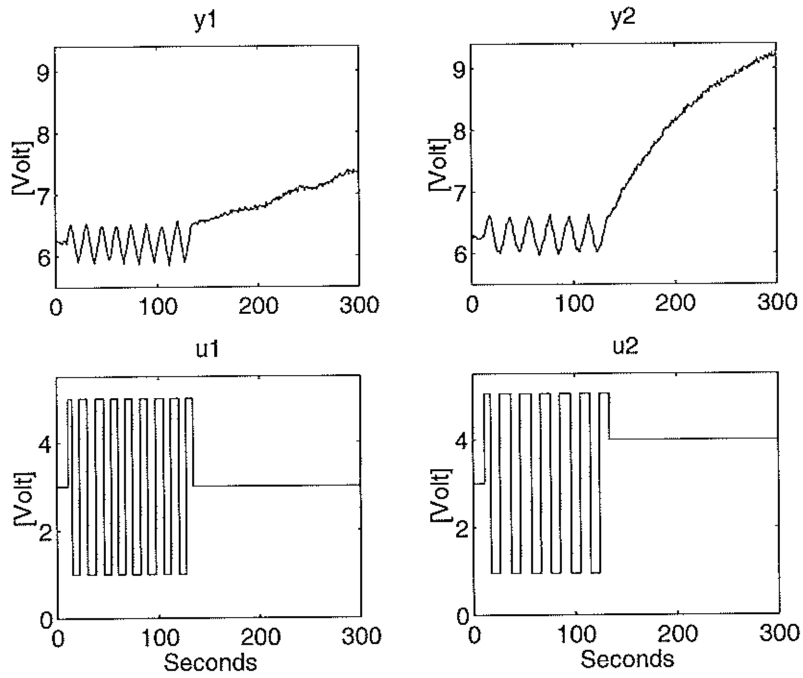

Figure 37 The second test applied to the minimum phase system.

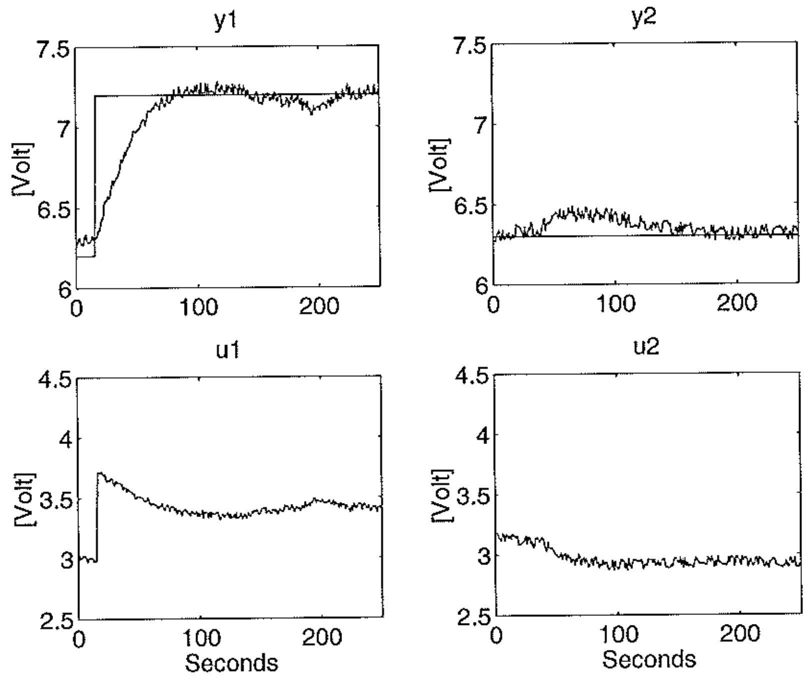

Figure 38 The results of the minimum phase system controlled by the proposed method.

where $\omega_c = 0.5(\omega_c^1 + \omega_c^2) = 0.208$. They are quite accurate compared with their "true" values:

$$G(0) = \begin{bmatrix} 1.5 & 2.5 \\ 2.5 & 1.6 \end{bmatrix}$$

$$G(j\omega_c) = \begin{bmatrix} 0.11e^{-1.52j} & 0.024e^{-2.96j} \\ 0.011e^{-3.01j} & 0.085e^{-1.52j} \end{bmatrix}$$

Unlike the minimum phase system, the outputs have the same oscillation frequencies. This is so because the interaction is larger in this case.

The simple identified model is used to design the controller in the same way as before.

The gain margin and phase margin are chosen as $A_{m1} = A_{m2} = 3$ and $\phi_{m1} = \phi_{m2} = \frac{\pi}{3}$. Then yield the following controller

$$C(s) = \begin{bmatrix} -0.43(1 + \frac{1}{83.3s}) & -0.019(1 - \frac{1}{2.28s} - 0.44s) \\ -0.0119(1 - \frac{1}{1.46s} - 1.19s) & -0.48(1 + \frac{1}{95.1s}) \end{bmatrix}$$

The tuning processes and the performance of the controller are not shown, but instead a closed loop stability analysis has been done.

In general a MIMO system is stable if all its poles are strictly inside the stability region (the LHP for continuous time system, and the unit disk for discrete time systems).

The closed loop control design is shown in Figure 35. The transfer function is given by

$$y = (I + G(s)C(s))^{-1}G(s)C(s)R(S)$$

where I is the identity matrix.

Then, the poles of the closed loop system are

$$P = \det(I + G(s)C(s))$$

The Figure 39 shows the poles and zeros of the closed-loop system. The pole-zero plot shows that this new controller design method doesn't work for our system. We see that we have poles in the RHP, i.e., the closed loop system is unstable.

Non-Minimum Phase System, control structure two We proceed in the same way as before and perform the experiments for the non minimum phase system, control structure two.

The relay in loop 2 is an ideal one with switching levels 0.2 and −0.2, and switching levels of the relay in loop 1 are 0.26 and −0.20 in the first test, and are then changed to 0.29 and −0.22 in the second.

In the first test the system exhibits limit-cycle oscillations of common frequency with frequency $\omega_c^1 = 0.0217$.

In the second test it exhibits a limit-cycle oscillations of common frequency, with $\omega_c^2 = 0.0215$.

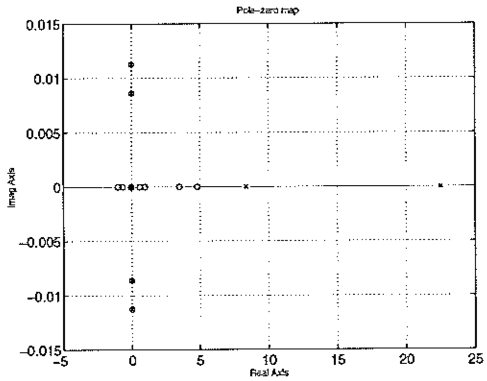

Figure 39 The poles and zeros of the closed loop system.

The steady-state gain matrix $\hat{G}(0)$ and the frequency-response matrix $\hat{G}(j\omega_c)$ for the this system are then computed. The results are

$$\hat{G}(0) = \begin{bmatrix} 3.3 & 1.6 \\ 1.7 & 3.3 \end{bmatrix}$$

$$\hat{G}(j\omega_c) = \begin{bmatrix} 1.33e^{-2.01j} & 0.90e^{-1.31j} \\ 0.78e^{-1.30j} & 0.90e^{-2.17j} \end{bmatrix}$$

where $\omega_c = 0.5(\omega_c^1 + \omega_c^2) = 0.0216$. They are quite accurate compared with their "true" values:

$$G(0) = \begin{bmatrix} 2.5 & 1.5 \\ 1.6 & 2.5 \end{bmatrix}$$

$$G(j\omega_c) = \begin{bmatrix} 1.12e^{-1.64j} & 0.89e^{-0.94j} \\ 0.72e^{-1.10j} & 0.72e^{-1.98j} \end{bmatrix}$$

We use the information that we have from the identification to design the controller.

The gain margin and phase margin are chosen as $A_{m1} = A_{m2} = 2$ and $\phi_{m1} = \phi_{m2} = \frac{\pi}{4}$. Then yield the following controller

$$C(s) = \begin{bmatrix} 0.03(1 + \frac{1}{26.7s}) & -0.06(1 + \frac{1}{78.9s} + 33.2s) \\ -0.05(1 + \frac{1}{68.6s} + 22.9s) & 0.07(1 + \frac{1}{63.7s}) \end{bmatrix}$$

The tuning processes are shown in Figure 40 and Figure 41. The performance of the controller is shown in Figure 42.

3.6 Summary

In this Section we presented a detailed study of various methods for automatic tuning of PID controllers for MIMO-systems.

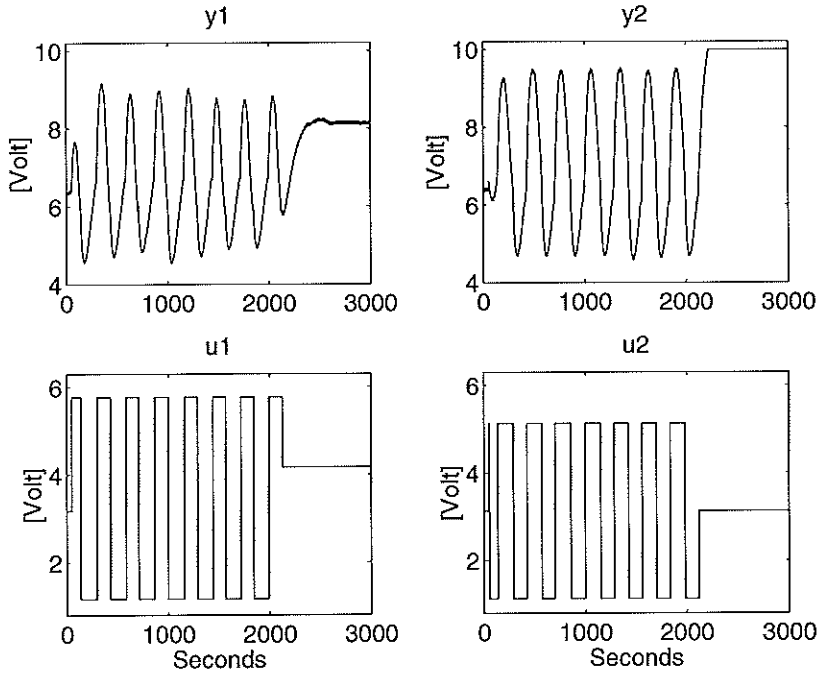

Figure 40 The first test applied to the non minimum phase system, control structure two.

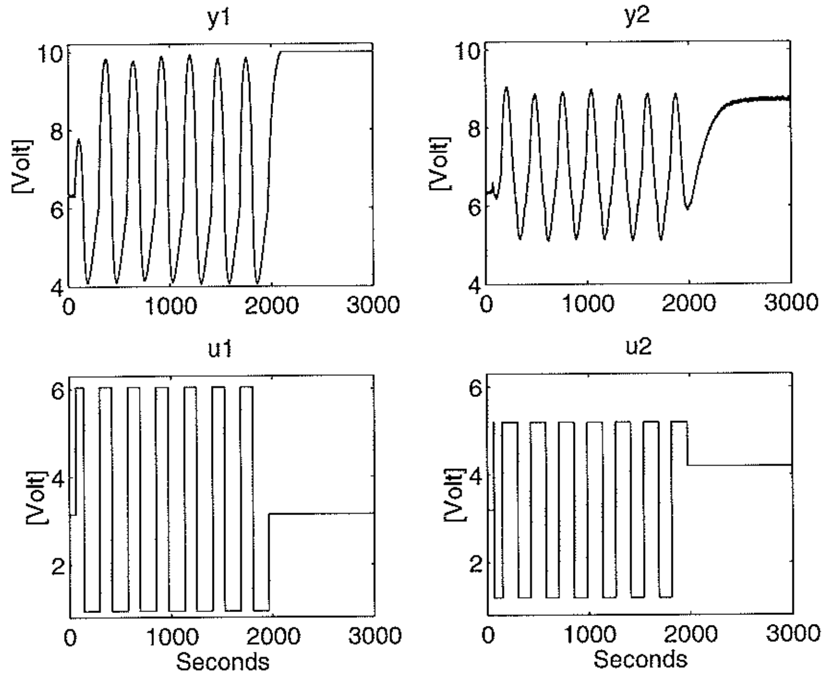

Figure 41 The second test applied to the non minimum phase system, control structure two.

With the help of Figure 43, Figure 44 and Figure 45 we give a summary of the closed-loop performance cfor the various methods. Note that the number of the collected points is not the same for all the experiments. The reason for

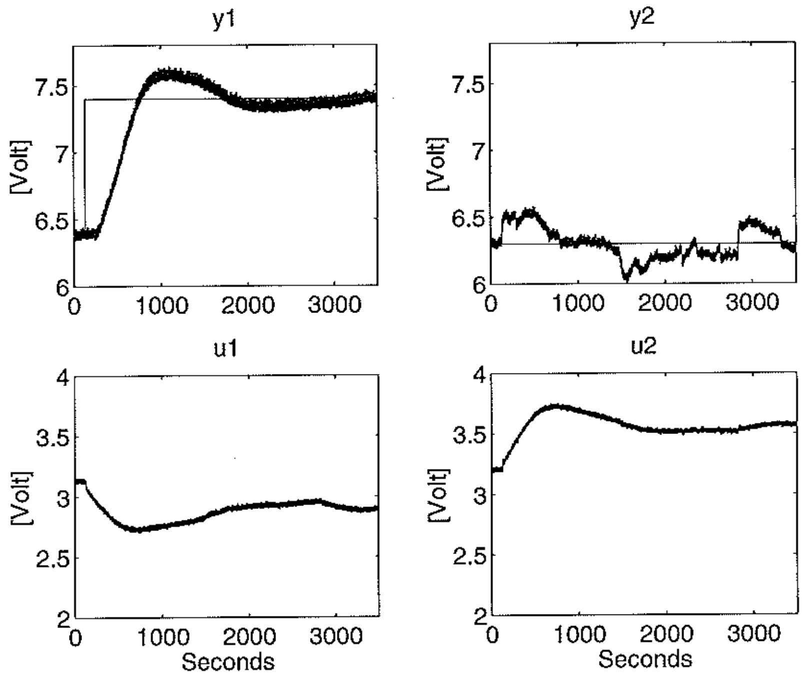

Figure 42 The results of the non minimum phase system, control structure two, controlled by the proposed method.

this is that for some method, the settling time was shorter and we stopped the experiment when the system was stabilized. For some other method the settling time was longer and thus the number of the collected points is larger. This is the reason for the existing gap on some of the plots.

The experiments show that the SISO-method compares favorably with the Sequential-method. The settling times for set point responses are significantly better for the SISO- and Sequential-method than for the decentralized-method.

It must be noted that the settling time is approximately ten times longer for the non minimum phase system. In Figure 45, the level of the interactions is lower in the decentralized-method than that achieved in the previous methods, but the settling time is significantly slower for this method.

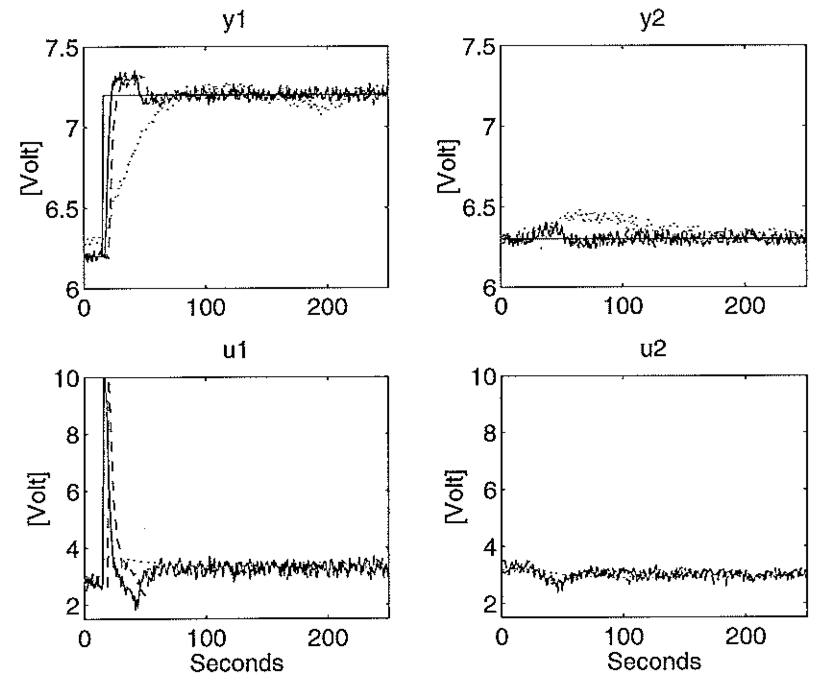

Figure 43 The results of the minimum phase system. The system was controlled by a controller designed as in the SISO-method (solid line), as in the Sequential-method (doted) and as in the decentralized-method (colons).

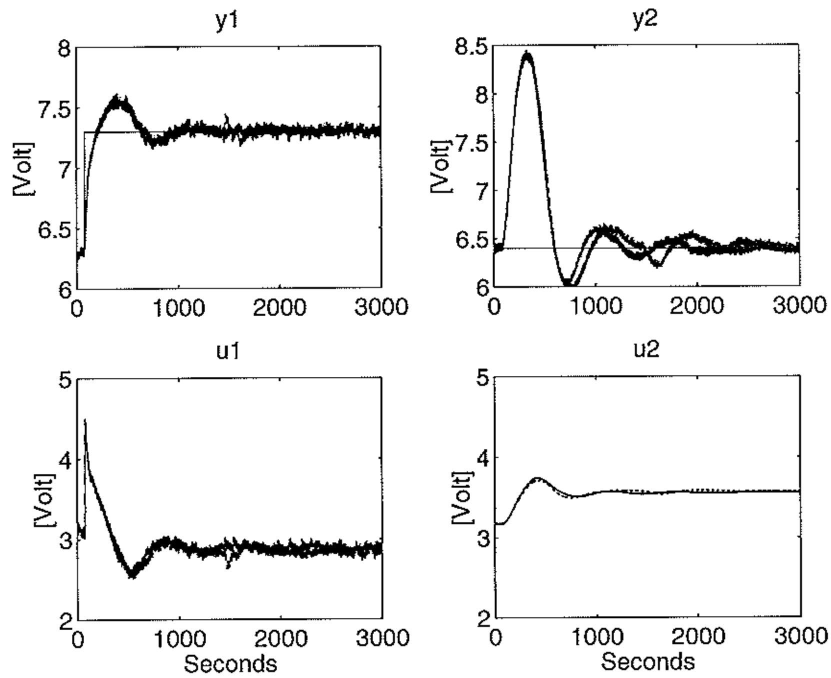

Figure 44 The results of the non minimum phase system, control structure one. The system was controlled by a controller designed as in the SISO-method (solid line), as in the Sequential-method (doted).

Moreover, the controller parameters designed according to the new derived

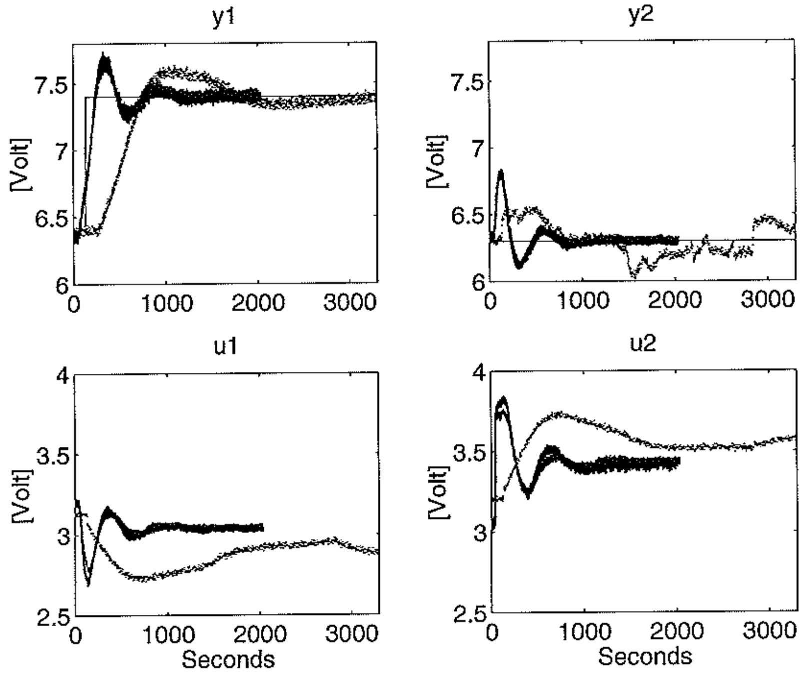

Figure 45 The results of the non minimum phase system, control structure two. The system was controlled by a controller designed as in the SISO-method (solid line), as in the Sequential-method (doted) and as in the decentralized-method (colons).

equations, gave reasonable results only in two cases, for the minimum phase system and for the non minimum phase system, control structure two. For the non minimum phase system, control structure one, the derived design equations didn't manage to stabilize the system at all.

So, the benefits of the decentralized method, which is a method much more complex than the previous methods, are negligible.

4. Implementation

4.1 Introduction

The complete system, including both the hardware and software is shown in Figure 46.

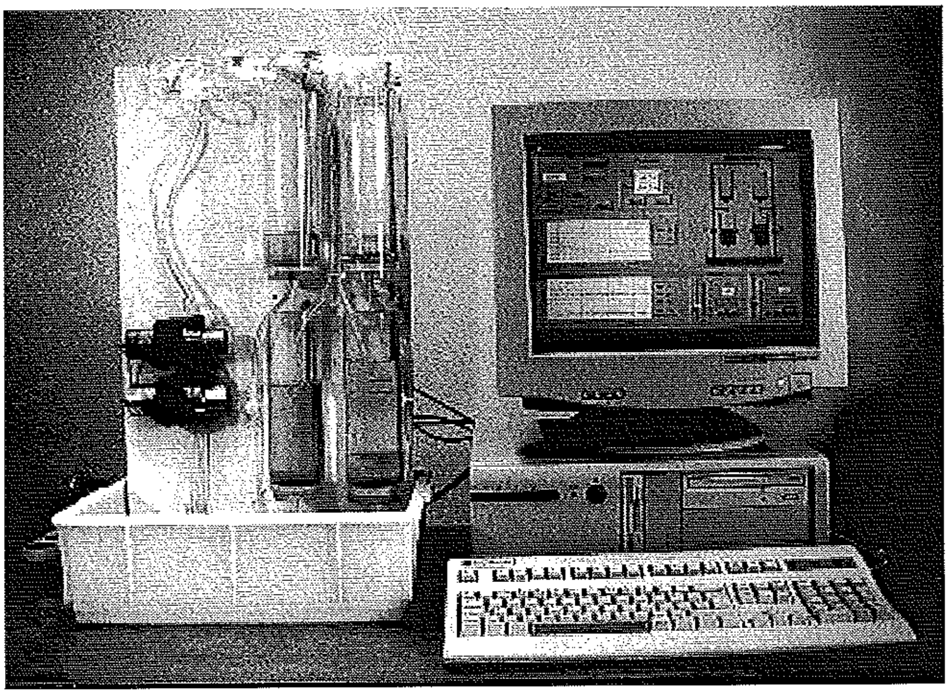

Figure 46 The quad tank laboratory system shown together with the controller interface running on a Pentium PC.

This Section starts with a short introduction to the software used for adding new features to the graphical interface built by another graduate student [Nunes, 1997].

The software used for developing the graphical interface was InTouch, which is a software package used to create PC based man-machine interfaces. The package consists of two major elements: WindowMaker and WindowViewer.

WindowMaker is the development environment, where object oriented graphics are used to create animated touch-sensitive display windows.

WindowViewer is the runtime environment used to display the graphic windows created in WindowMaker.

An InTouch application is built up interactively by drawing an interface, filling in forms and writing scripts. We are given the possibility to bring life to a graphic object or symbol by attaching Animation Links to it, which makes the object to change appearance, to reflect changes in the value of a variable or an expression. In WindowMaker the Tagname Data Dictionary is available. This is a runtime database containing the current value of all the variables used. In order to create the runtime database, InTouch requires information about all the variables being created [Årzén, 1996].

Interprocess communication in Windows is handled by DDE, which is the acronym for Dynamic Data Exchange. DDE is a message based protocol that implements a client-server relationship between two applications.

All the algorithms are implemented in Modula2, which runs over the Real-Time Kernel. The program is structured in modules. There are four processes:

1. OpCom which handles the communication with the InTouch program.
2. RefGen that generates the reference signal waves.
3. Regul which communicates with the real process.
4. Main which takes care of starting and terminating the processes

4.2 User's Guide

In this Section we explain how to use the interface (only the parts that we have added), for more details see [Nunes, 1997].

The machine was connected with the quad tank system through an AD/DA converter with four inputs and two outputs.

The interface consists of two screens, Main Screen and Setup Screen. In Main Screen we have different types of modes. Here follows a very short description:

Simul Simulations on the non-linear model.

Active Runs on the real system.

Setup Change parameters that are not changed so often.

Man Produces a control signal to the system by increasing or decreasing the voltages to the pumps manually.

Controller Controlling the system by different controllers

Tuning Different relay experiments. This mode together with what belongs to it is completely new.

A more detailed explanation of what we can do in Tuning Mode follows here. When we click on the Tuning button we get the Relay box-window. Inside it we see the type of the relay experiment that have been chosen. Note that, the first time we click on the Tuning button we see the default type with its default parameters. Then, if we decide to do another relay experiment, we can click on the Type button and choose one of the alternatives. See Figure 47.

If we want to change the relay parameters we just have to click in that particular Relay in the Relay matrix, and then automatically appears a small window. See Figure 48. Then its possible to change the parameters. All parameters are changed by using numerical inputs. By clicking at the check box we will activate an integrator in series with the relay. The relay parameters are:

- d the relay amplitude.
- ϵ the hysteresis.

In the Relay Controller box there is also an extra feature that we mentioned before and that is the integrator. If we activate it, it means that we have an integrator in series with the relay.

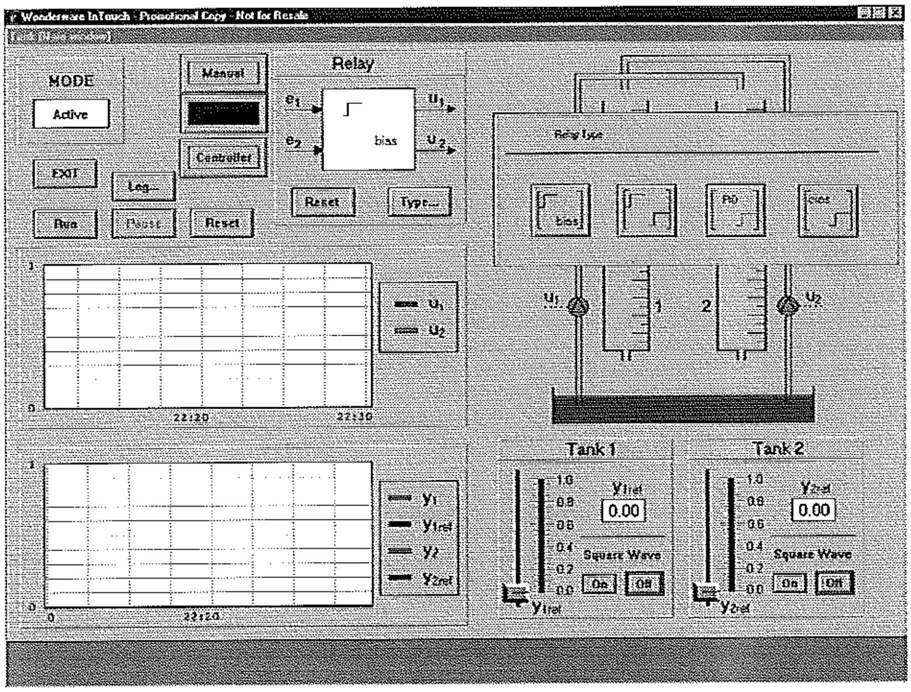

Figure 47 The relay box.

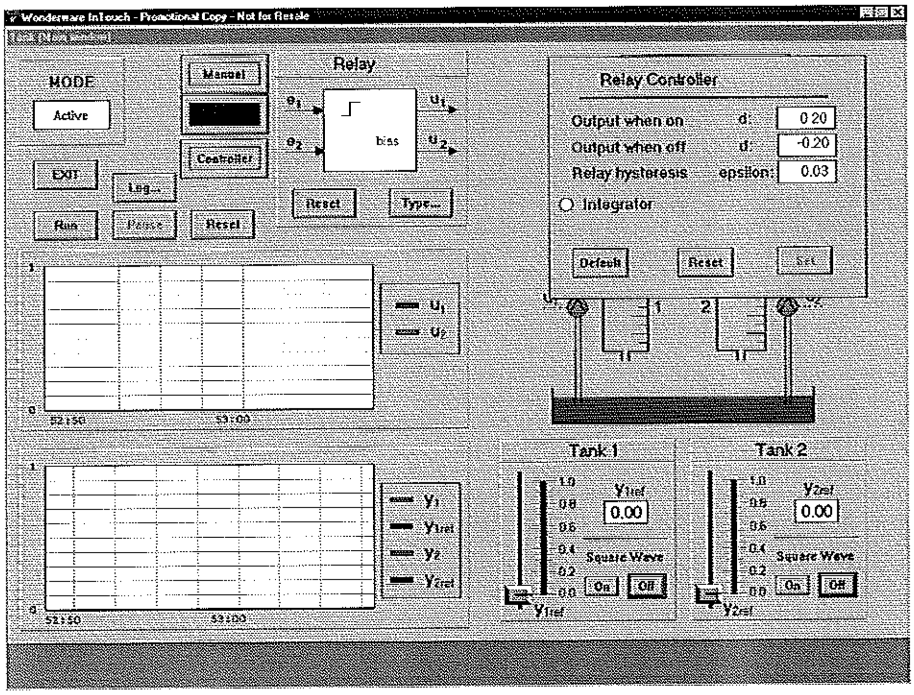

Figure 48 The relay parameters.

When we did our relay experiment for the identification of simple models, we moved the system manually to a steady state around a working point. Then we switched to the tuning mode. See Figure 49.

In the Setup Mode the user can change parameters that are not changed so often, as for example the sampling time, simulations parameters etc. Here you

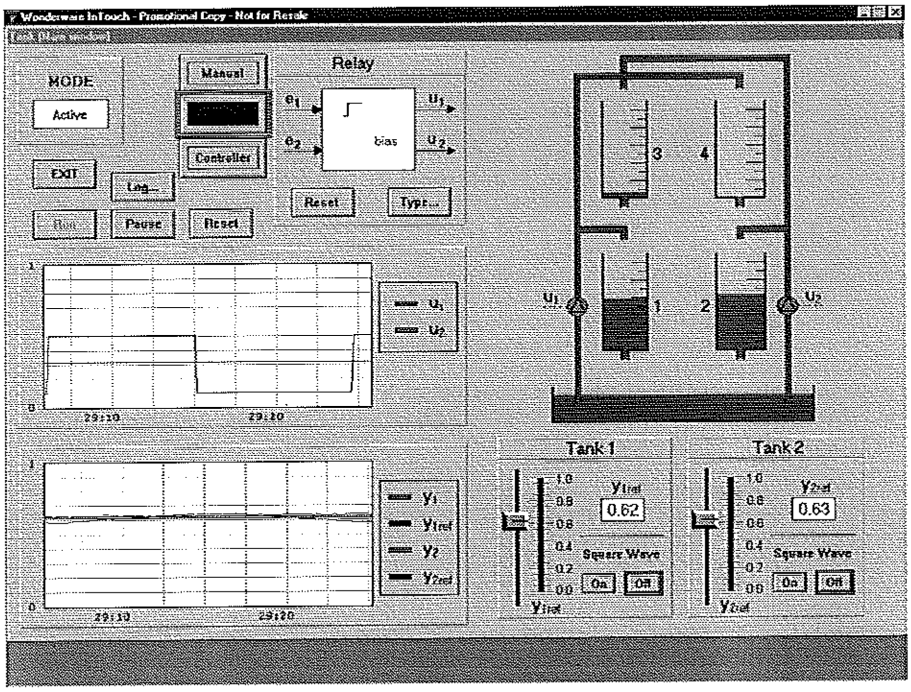

Figure 49 A SISO-relay experiment.

can also see a block diagram of the system, i.e., how the system inputs and outputs are connected to each other, what kind of controller we have (relay, PID or we control it manually). See Figure 50.

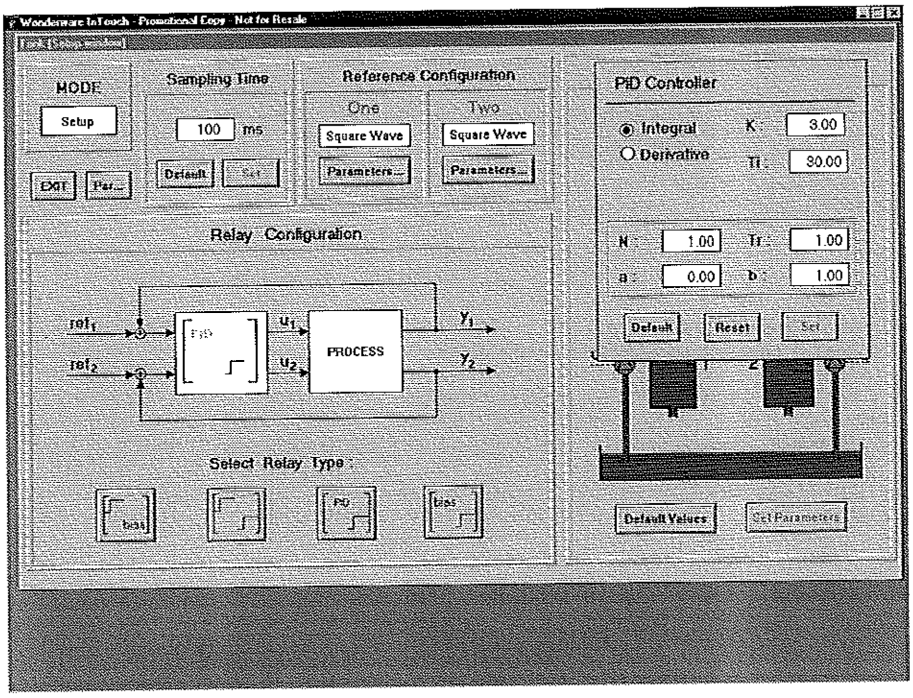

Figure 50 The Setup Mode.

5. Conclusions

In this thesis a multivariable laboratory control system has been described. The derived physical linear model of the system has a multivariable zero, the position of which could be varied by changing the position of two valves. So by doing that our system was easier or harder to control.

Auto-tuning of controller requires some information on the process dynamics, and this may be obtained by injecting a test signal into the process. Åström and Hägglund proposed a relay feedback auto-tuning technique that can approximately determine the critical gain and critical frequency of the process [Åström and Hägglund, 1984]. This has been shown to be very efficient for SISO-systems.

In this thesis, this idea is extended to our multivariable system. This extension has been investigated using various methods. These are:

SISO Controller Tuning Only one loop at at a time is subjected to relay feedback, while the second loop is kept open. The controller for the first loop is designed. Then, the second loop is subjected to relay feedback while the first loop is kept open. The controller for the second loop is designed.

Sequential Controller Tuning A loop is closed with a simple controller once a relay test had been done to that loop. The first controller is constructed based on one sub-process while all the other sub-processes remain on open loop. The next controller is constructed with the previous designed loop closed.

Extension to Sequential Controller Tuning The sequential method is extended to identify a more accurate model and in that way to design a more efficient controller. This method is based on the observation that, a system with a relay in a series with a filter, will give any point on the Nyquist curve.

Decentralized Controller Tuning This is a new method for auto-tuning of decentralized PID controllers [Wang et al., 1997]. In the tuning mode all the controllers were replaced by relays, and a critical point was identified from the limit cycles reached in the two loops via the derived relations. An algorithm to obtain the response at the desired critical point was presented. Steady-state gains were identified from open-loop step tests which is different from the method in [Wang et al., 1997]. This was still achieved within the two performed experiments, but it took longer time to complete the identification.

The sequential method is often better than the SISO method. It offers several advantages. It is very simple and stability is ensured at every stage of the design through sequential loop closing. Moreover, the sequence of loop closing is important since it affects the amount of interaction entering all the previously tuned loops and therefore limits the quality of the set point responses for those loops. It is generally expected that when the faster loop is tuned first, it has the advantage that it is less affected by interactions. More importantly, it allows the slower loop to be tuned last and hence be able to account for the interactions resulting from the closure of the faster loop [Vasnani, 1994]. But for our plant, where the critical frequencies of the diagonal elements are similar, it has been shown through experiments that the sequence of tuning is not so significant.

The extended sequential method doesn't manage to design a better controller than the SISO-, or sequential method. This is so, because the two identified points (point 1 and point 2) are very close to each other. Thus the model achieved in this way is not much more complex than the model identified in the previous methods.

The decentralized method doesn't work when the interaction is large or when the relay experiments give oscillations with different frequencies.

The controller parameters designed according to the new derived equations, gave reasonable results in two cases. It seems that, the derived design equations, don't manage to design a controller for a system with large interactions. Furthermore, when the interactions were small, the identification procedure didn't manage to give a model with good accuracy.

All the experiments in this thesis have been performed using a PC interface that has been developed in the man-machine interface generator InTouch.

Acknowledgments I am very grateful to Per Hagander who gave me the opportunity to do this master thesis.

I would like to thank my supervisor Karl Henrik Johansson for his continuous help and encouragement during my work.

6. References

[1] Arzén, Karl-Erik. *Real-Time Control Systems.* Departement of Automatic Control, Lund Institute of Technology, Sweden, 1996.

[2] Astor, Eric. *Från Pascal till MODULA-2.* Studentlitteratur, Lund, Sweden.

[3] Åström, K. J. and T. Hägglund. *PID Control Theory, Design & tuning.* Departement of Automatic Control, Lund Institute of Technology.

[4] Åström, K. J. and T. Hägglund. *Automatic Tuning of Simple Regulators with Specifications on Phase and Amplitude Margins.* Automatica 20, Special issue on adaptive control, 1984.

[5] Bristol, E. *On a new meassure of interaction for multivariable process control.* IEEE Transactions on Automatic Control, 11, p. 133, 1966.

[6] Holt, B. R. and Manfred Morari. *Design of Resilient Processings Plants. The Effect of Right-Half-Plane Zeros on Dynamic resilience.* Chemical Engineering Science, Vol40, 1985.

[7] Hägglund, T. and Åström, K. J. *Industrial Adaptive Controllers Based on Frequency Response Techniques.* Automatica, Vol. 27, No. 4, pp. 599-609, 1991.

[8] *Intouch - Users Guide.* Wonderware Corporation, 1995.

[9] Johansson, K. H. *Relay Feedback and Multivariable Control.* PhD thesis, Departement of Automatic Control, Lund Institute of Technology, 1997.

[10] Morari, M and E Zafiriou. *Robust Process Control.* Prentice Hall, 1989.

[11] Niederlinski, A *A Heurestic Approach to the Design of Linear Multivariable Interacting Control Systems.* Automatica, Vol. 7, 1971.

[12] Nunes, José Luís. *Modeling and Control of The Quadruple-Tank Process.* Master Thesis, Departement of Automatic Control, 1997.

[13] Wang, Qing-Guo, Biao Zou, Tong-Heng Lee and Qiang Bi. *Auto-tuning of Multivariable PID Controllers from Decentralized Relay Feedback.* Automatica, Vol.33, 1997.

[14] Vasnani, V. U. *Towards Relay Feedback Auto-Tuning of Multi-Loop Systems..* PhD thesis, National University of Singapore , 1994.

[15] Ziegler, J. G. and N. B. Nichols. *Optimum settings for automatic controllers.* Trans. ASME, 65, p. 433-444, 1943.