Ramkrishna Ghosh

# Input Designs for Identification of Ill-conditioned Multivariable Systems

**Ramkrishna Ghosh**

- M.Sc. in Instrumentation Science, 2006, University of Pune, India.

- Worked as Design and development Engineer, Forbes Marshall Ltd., India (2006-08).

- Research Fellow at CAD Centre, Chemical Engineering, I I T Bombay, India (2008-09).

Since Sept. 2009, researcher at Åbo Akademi University, specialized on process control, system identification and process data-based alarm management.

**Supervisor**

Professor Kurt-Erik Häggblom

Åbo Akademi University, Turku

**Opponent and Reviewer**

Professor Sirkka-Liisa Jämsä-Jounela

Aalto University, Espoo/Helsinki

**Reviewer**

Doctor Mats Friman

Metso Flow Control Oy, Tampere

# Input Designs for Identification of Ill-conditioned Multivariable Systems

Ramkrishna Ghosh



Ph.D. Thesis in Process Control
Faculty of Science and Engineering
Åbo Akademi University

Åbo/Turku, Finland 2016

*Dedicated to*

***My Parents***
*Shri. Ramesh Chandra Ghosh (Baba),*
*Smt. Minati Ghosh (Ma),*
*for their constant support and encouragement*

*&*

*to my beloved **Teachers**,*
*who have cultivated me to reach here.*

# Preface

$\text{T}$he work in this thesis was carried out in the process control laboratory at Åbo Akademi University. First, I would like to express my gratitude to my advisor, Prof. Kurt-Erik Häggblom, for his guidance, motivation and consistent encouragement I received throughout the research work and writing of this thesis. I consider it as a great opportunity to learn from his research expertise. I am grateful to him for the help and support. I would like to thank Dr. Jari Böling for his guidance, encouragement and insightful comments during my research work. My gratitude goes to Professors Hannu Toivonen, Mikko Hupa, Henrik Saxén, Mani Bhushan, Sachin Patwardhan, Sharad Bhartiya, Dulal Sen, Amarendranath Chatterjee and Doctors Tore Gustafsson, Frank Petterson, Anirban Neogi, Basudev Ghosh, Abhay Bulsari and other teachers for their helpful career advices and encouragements.

My gratitude goes to Prof. Johan Schoukens from Vrije Universiteit Brussel (VUB) for providing me the opportunity to take part in VUB spring school and funding my stay in Brussels. It was a great learning experience to share with fellow international researchers in the field of system identification and control. A special thanks goes to Prof. Henrik Saxén for giving the opportunity to work with him on a collaboration project related to blast furnace with researchers from Zhejiang University, China. This is very special because during this time I was in need of financial support for continuing my thesis.

# Abstract

Identification of an industrial system is generally a costly and complex process. Making an identification procedure efficient enough is an ongoing demand. Typically, designing the excitation for ill-conditioned and directional systems is a challenging task. Tailor-made input excitation using process knowledge is one way to make the identification more efficient. The focus of this thesis is to develop a better identification process for multiple-input multiple-output (MIMO) systems with proper design of input excitation signals. A distillation column simulator has been constructed as a testbed for the research. Data from a real-life distillation column system have been used for calibrating the simulator.

In model-based control (e.g., model predictive control, MPC), the model quality has a critical effect on the performance of the controller. To construct a control-relevant model via identification, the system has to be excited adequately in all gain directions. In order to reduce the relative uncertainties associated with these gain directions, it is especially important to excite the system in the weak gain directions. In this study, we consider both designs by rotated signals as well as methods based on design in the frequency domain. We study multiple input design methods using basic signals like steps, PRBS, and sinusoidal signals with multiple frequencies and compare them through two case studies. The case studies are a $4 \times 4$ column stripper system and a $2 \times 2$ nonlinear distillation column system. As most of the previous studies have focused on $2 \times 2$ systems, a comparative study to choose the proper excitation for identification of an ill-conditioned system with more inputs and outputs ($n \times n, n > 2$) was chosen.

Practically every system has physical limitations as well as limitations due to the operation of the process. A useful measure related to input limitations is the plant-friendliness of an input signal. Several parameters that characterize the plant-friendliness are calculated and compared for the designed inputs. These parameters depend only on the inputs, not the rest of the system.

For the model identification, it is desirable that outputs are excited equally in all directions. We propose two different tools for analyzing the excitation of the outputs, namely, projections of the outputs along the gain directions and the determinant of the correlation matrix of the outputs. Both methods yield measures on how well-balanced the output distribution is. We show that experiments where the gain directions have been considered in the input design produce output data that are better balanced than output data produced by other designs. The models obtained from directional experiments perform better in cross-validation with data from other experiments than other models.

# Svensk sammanfattning

Identifiering av ett industriellt system är allmänt en dyr och komplex process. Det finns ett fortgående behov av tillräckligt effektiva identifieringsmetoder. Utformning av exciteringen för illakonditionerade och riktningsberoende system är typiskt en utmanande uppgift. Skräddarsydd excitering av insignalerna genom utnyttjande av processkunskap är ett sätt att göra identifieringen effektivare. Fokus i denna avhandling är utvecklandet av en bättre identifieringsmetodik för system med multipla insignaler och multipla utsignaler (MIMO) genom en ändamålsenlig utformning ("design") av insignalexciteringen. En simulator för en destillationskolonn har konstruerats och använts som testbädd i forskningen. Data från en verklig destillationskolonn har använts för kalibrering av simulatorn.

I modellbaserad reglering (t.ex. modellprediktiv reglering, MPC) har modellkvaliteten en avgörande betydelse för regulatorns prestanda. För att erhålla en reglerrelevant modell genom identifiering bör systemet exciteras lämpligt i alla förstärkningsriktningar. Det är speciellt viktigt att excitera systemet i riktningar med svag förstärkning för att reducera de relativa osäkerheterna förknippade med förstärkningsriktningarna. I denna studie beaktar vi design med roterade signaler såväl som metoder baserade på design i frekvensplanet. Vi studerar designmetoder för multipla insignaler genom användning av elementära signaler såsom steg, PRBS och sinussignaler med multipla frekvenser och jämför dem i två fallstudier. Fallstudierna är ett kolonnstrippersystem av storleken $4 \times 4$ och ett olinjärt destillationskolonnsystem av storleken $2 \times 2$. Eftersom de flesta tidigare designstudierna gäller $2 \times 2$-system, valdes en jämförande studie av lämplig excitering för identifiering av ett illakonditionerat system med fler insignaler och utsignaler ($n \times n, n > 2$).

Varje process har i praktiken såväl fysikaliska begränsningar som av driften föranledda begränsningar. Ett nyttigt mått förknippat med begränsningar på insignalerna är insignalernas processvänlighet (eng. plant-friendliness). Flera parametrar som karaktäriserar processvänligheten beräknas för de designade insignalerna och jämförs sinsemellan.

Dessa parametrar beror enbart av insignalerna, inte processen i övrigt. För modellidenti-fieringen är det önskvärt att utsignalerna exciteras lika kraftigt i alla riktningar. Vi föreslår två metoder för analys av utsignalexciteringen, nämligen projektioner av utsignalerna på förstärkningsriktningarna och determinanten av utsignalernas korrelationsmatris. Båda metoderna ger mått på hur väl balanserad utsignalspridningen är. Vi visar att experiment där förstärkningsriktningarna har beaktats i signaldesignen ger en bättre balanserad utsignalspridning än andra experiment. Modellerna från riktningsbeaktande experiment fungerar bättre i korsvalidering med data från andra experiment än andra modeller.

# Suomenkielinen yhteenveto

Teollisten järjestelmien identifiointi on yleensä kallis ja monimutkainen prosessi. Häiriöille ja tulosignaalien yhteisvaikutukselle herkkien järjestelmien identifiointi on erityisen haastavaa. Räätälöity tulosignaalien suunnittelu prosessituntemuksen avulla on eräs tapa tehostaa identifiointia. Tämän väitöstyön keskeisenä aiheena on kehittää parempi identifiointiprosessi monitulo-monilähtö (multiple-input and multiple-output, MIMO) -järjestelmille suunnittelemalla tulosignaalien asianmukainen viritys.

Mallipohjaisessa prosessinohjauksessa (model predictive control, MPC) käytetyn mallin laadulla on ratkaiseva vaikutus saavutettuun suorituskykyyn. Hyvä malli kuvaa prosessia tarkasti kaikilla tulosignaalien kombinaatioilla. Tällainen malli voidaan rakentaa järjestelmäidentifioinnin avulla, kun järjestelmää kuormitetaan riittävästi kaikkiin suuntiin. Suhteellisten epävarmuuksien vähentämiseksi on erityisen tärkeää kuormittaa järjestelmää heikkojen vahvistuksien suuntiin.

Tulosignaalien suunnittelussa käytetään perusaskelsignaaleja, PRBS-signaaleja, sekä sinimuotoisia signaaleja useilla taajuuksilla. Näitä verrataan tässä työssä keskenään kahden eri sovelluksen avulla; toinen sovellus on lineaarinen erotuskolonnijärjestelmä (stripperi) neljällä tulo- ja lähtösuureella ($4 \times 4$–järjestelmä), toinen on epälineaarinen tislauskolonnijärjestelmä kahdella lähtö- ja tulosuureella ($2 \times 2$– järjestelmä). Useimmat aiemmat tutkimukset ovat keskittyneet lineaarisiin $2 \times 2$– järjestelmiin.

Työssä on kehitetty kaksi erillistä työkalua järjestelmän virityksen tutkimiseksi; lähtösignaalien heijastaminen eri vahvistuksien suuntiin sekä niiden korrelaatiomatriisin determinantin analysointi. Molemmat työkalut tuottavat tietoa lähtösignaalijakauman tasapainoisuudesta. Kokeet, joissa vahvistuksien suunnat on otettu huomioon tulosignaalien suunnittelussa tuottavat lähtösignaalidataa, joka on paremmin tasapainossa kuin muissa kokeissa. Suuntauskokeista saadut mallit toimivat ristiinvalidoinnissa muiden koetulosten kanssa paremmin kuin muut mallit.

# Contents

# List of Figures

# List of Tables

# Introduction

With the advancement of technology, systems are becoming more complex. In order to achieve better and efficient performance, we are continuously pushing the boundary where a system can be operated. The demand to improve the outcome, and to achieving close to perfect operation of a system, often leads to these underlying questions:

1. How well can we **model** a true system?

2. If a model is not available, what is the best possible way to **identify** the model with minimum effort?

3. What is the best possible **excitation** one should employ to minimize the identification cost?

4. Which criteria (e.g. cost function and constrains) should be used to obtain desired performance in the **optimization** process?

5. What optimization **algorithm** should be used?

Interestingly, these are quite common and active research areas in various domains of engineering and science. Often, multidisciplinary skills are required to tackle these issues.

An accurate and reliable model of a system is vital not only in engineering applications, but also in biology, bio-informatics and many more areas. The purpose of the model in all of these depends on the application. For example, in biology a model might be employed to predict spread of a disease and to prevent or minimize the casualty. On the other hand,

in case of process engineering a model is often used in model predictive control (MPC) or for tuning a feedback controller of a system to accomplish better performance.

Identifying a multivariable system is an active research area. Designing better excitations of the underlying system and making an identification procedure more informative are important tasks. Designing excitation for ill-conditioned and directional systems is often a challenging task. The focus of this thesis is to develop a better identification process with proper design of input excitation signals. A distillation column simulator has been constructed as a testbed for the research. Data from a real-life distillation column system in our laboratory have been used for tuning the simulator in order to achieve practically comparable results.

## 1.1  Thesis outline

The thesis is divided into seven chapters (1–7). In Chapter 2, we give a brief introduction to mathematical modeling and system identification in particular. Chapter 3 reviews some well-established identification material, which provides the theoretical framework for design of input signals. Chapter 4 is devoted to a discussion on plant friendly identification procedures and multiple related parameters. In Chapter 5, we discuss and analyze existing methods and their applications in tackling ill-conditioned system identification in literature. In Chapter 6, we discuss our proposed method, experiments and results with two real world examples, a distillation column system (a $2 \times 2$, 2 inputs-2 outputs ill-conditioned system) and another industrial distillation column system (a $4 \times 4$, 4 inputs-4 outputs ill-conditioned system) . A comparative study of different methods has been performed on both case studies. Research outcomes are concluded in Chapter 7.

## 1.2  Main contributions

Most of the basic input excitation signals (e.g. step, PRBS and multisine) are well-known in system identification literature. However, a comparative study to choose the proper excitation to identify a higher order ($n{\times}n$ with $n > 2$) inputs-outputs ill-conditioned system is new. Two different tools are proposed for analyzing the excitation of the outputs. Both methods yield measures on how well-balanced the output distribution is.

## 1.3 Research publications

The work presented in this thesis is based on the following peer-referred journal/ articles and peer-referred conference proceedings. The contributions of the thesis author in the publications is also elucidated.

**Paper I** R. Ghosh, K. Häggblom, J. Böling. Control-relevant input excitation for system identification of ill–conditioned $n \times n$ systems with $n > 2$. *19th IFAC World Congress*, pp. 9382-9387, Cape Town, South Africa. August 2014.

Author's contributions: The thesis author participated in the problem formulation and experiment design, did the simulations and the modeling, wrote the first draft manuscript and finalized the paper together with the co-authors.

**Paper II** R. Ghosh. Construction of plant–friendly input excitations for ill–conditioned multivariable process identification. *IEEE International Symposium on Communications, Control, and Signal Processing (ISCCSP)*, Athens, Greece. May 2014.

Author's contributions: The thesis author did the main work and wrote the paper by himself.

**Paper III** R. Ghosh, K. Häggblom, J. Böling. Construction of plant-friendly input excitation for ill–conditioned multivariable process in open loop identification for $n \times n$ systems with $n > 2$. (*submitted*).

Author's contributions: The thesis author participated in the problem formulation and experiment design, did the simulations, modeling, and cross validations. He wrote the first draft manuscript and finalized the paper together with the co-authors.

**Paper IV** K. Häggblom, R. Ghosh. Experiment designs for control-oriented MIMO identification. *Finnish Society of Automation Publication Series no. 44*, Helsinki, Finland. March 2015.

Author's contributions: The thesis author participated in the problem formulation and design of experiments and contributed to the writing of the paper.

**Paper V** K. Häggblom, R. Ghosh. Tuning a distillation column simulator. *21st European Symposium on Computer Aided Process Engineering*, volume 29 of *Computer Aided Chemical Engineering,* pp. 823-827. Elsevier, 2011.

Author's contributions: The thesis author participated in the problem formulation, simulations, data analysis and model parameter fitting. He contributed to the writing of the paper.

Apart from the included publications, the thesis author made contributions to the following.

N. Mundru, S. Lalam, R. Ghosh, M. Bhushan. Cost optimal alarm design using data reconciliation. (*submitted*).

R. Ghosh, J. Böling, K. Häggblom. Balanced input excitation for identification of ill-conditioned $n \times n$ systems with $n > 2$. *Nordic Process Control Workshop* (NPCW) 2015, Troindem, Norway.

R. Ghosh, K. Häggblom, J. Böling. Designing balanced input signal for identifying ill-conditioned systems. *Nordic Process Control Workshop* (NPCW), 2013, Oulu, Finland.

R. Ghosh, K. Häggblom. Commissioning a distillation column simulator. *Nordic Process Control Workshop* (NPCW), 2010, Lund, Sweden.

Modeling and identification of blast-furnace Si %, with Prof. Henrik Saxén, Thermal and Flow Engineering Laboratory, Åbo Akademi (July - December, 2012) (in a joint research program with Zhejiang University, Hangzhou, China).

# 2

# Process Modeling

A model  commonly serves as a tool to understand and explain observations and an abstraction of reality. Typically a real life system is non-linear in character and quite complex to model completely. Often a model is designed to characterize certain operating areas or states instead of the true system as a whole. The modeler has to decide how complicated and detailed the model should be depending on its end use or purpose. Process models  can be built in numerous ways. Considering the vast area of applications, typically model can be made in different ways in practice, however there is no one specific approach which can be considered as superior to the rest. Some approaches are summarized in Figure 2.1. More detailed discussions on this topic can be found in Söderström and Stoica [1989], Ljung [1999], Ogunnaike and Ray [1994] and Cameron and Hangos [2001].
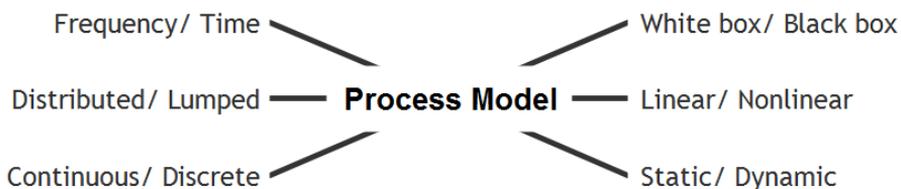
**Figure 2.1:** Different types of process models

Depending on the level of utilization of process knowledge or data, modeling can be divided in three major categories, namely a) Black box modeling, b) White box modeling and c) Gray box modeling; also summarized in Figure 2.2.

## 2.1   A theoretical modeling approach

In the case of white box modeling, models are constructed based on physical properties and prior knowledge of the system without using any input-output data. This approach exclusively use physical properties  for the model parameters  with all constant parameters used in the model known a priori.  The approach is often referred to as physical modeling or first principles modeling.

## 2.2   An experimental modeling approach

In black box modeling, the model is derived from input-output data of the system without considering any physical insights. Typically the model structure is chosen to be flexible and may be influenced by the success in other applications.  In practice they act as *variables* to fit the model with the observed data as well as possible. This implies that the model parameters  may or may not necessarily have any physical significance of the system.

## 2.3   A practical modeling approach (best of both)

In practice, white box and black box modeling are often combined together.  This kind of modeling is referred to as gray box modeling. Often, the model parameters are tuned by measured input-output data. Most real world systems comprise of elements such as reaction rate coefficients, friction coefficients, etc.. Determining their exact values in a system is close to impossible from the physics alone. Therefore, the above approach is more logical compared to the white box modeling.

## 2.4   Discussion

Black box modeling is occasionally used as almost synonyms to *system identification*; however this is not absolutely correct. "*System identification deals with the problem of building mathematical models of dynamical system based on observed data from the system.*" [Ljung, 1999], would be more correct definition and this is what we have used in most cases. Therefore, if we are using only white box model or white box model and tuning
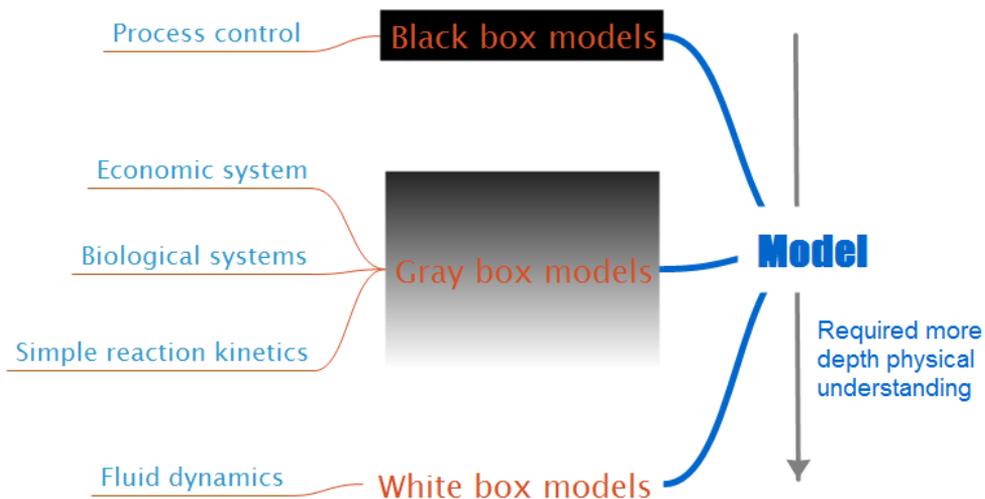
**Figure 2.2:** Schematic of modeling approaches

few system parameters (constants in the model equations) according to the observed data, we still consider it as system identification. This also means all identification procedures can be characterized in the scale ranging from pure white box and pure black box. White box and black box modeling are the two extreme cases though using one after another, less or more, is a common practice.

Let us now discuss all three approaches in detail with a real life example of a distillation column system. In Figure 2.3, a schematic of a typical distillation column system is shown. In this system, if we use the *first principles modeling* approach, we need to make a model for all small components and their physical properties. For this we need in depth knowledge of various constituents and their properties (physical, mechanical and chemical). As a result there would be hundreds of equations and variables, as used in our distillation column simulator , (more explanation can be found in Appendix A). Here, we need to consider all of its subsystems (e.g. trays, condenser, reboiler etc.) and model them using its physical properties in mathematical form. The following mathematical equations can be used for modeling a distillation column(Energy balances, Mass balances etc.) more details can be found in Appendix A.

Based on the actual size of a column there would be different number of subsystems,
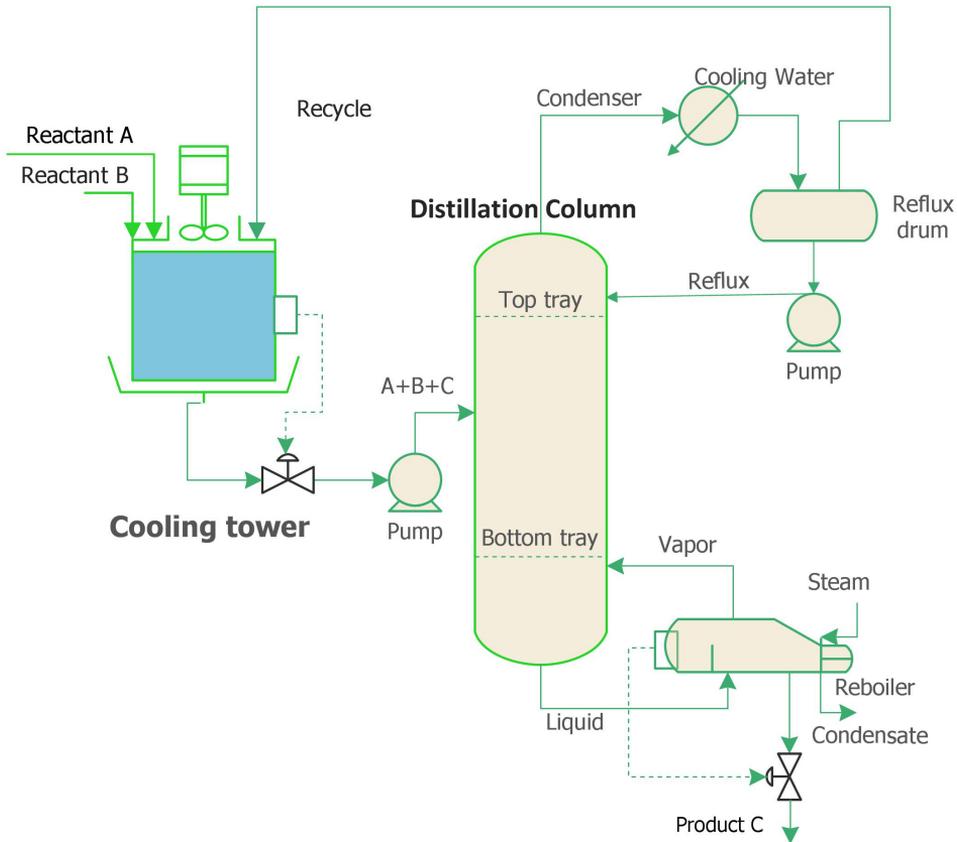
**Figure 2.3:** A schematic process diagram of a distillation column system

for example in our case, we have 15 trays along with the boiler and reflux. There are few equations and variables for each subsystem. Therefore, we need to solve $\approx 100$ equations (for 17 subsystems), as in Appendix A.

On the other hand, if we decided to use data based modeling (black box modeling), we can measure the input (reflux flow $L$ and vapor flow $V$) and output (top concentration $x_D$ and bottom concentration $y_B$) and make a model based on them. In this case the user does not need to have full system knowledge but only data, for the same example.

Defining parameters from physics of the system while tuning them with the help of input output data, would be a better approach (gray box modeling), as we have explained in Section 2.3. One practical advantage of this method is flexibility (modularity) in a simulator system (e.g. a subsystem model can easily be reused, for example in another

distillation column maybe with 50 trays and the parameters can be fine tuned with input output data from such a system). This method would be quite useful in an industrial scenario where typically many distillation columns are used in a chemical process.

The process to estimate all parameters is a big task, and often in an optimization finding a correct set of values with practical significance is difficult (e.g. time delays of a system might deviate considerably from the values suggested by an optimization algorithm). Moreover, it is much easier to handle a model with lower number of variables. However, a first principles model, which contains detailed information of physical properties, is better in treating various operating regions than a black box model, which is only valid under the conditions that excited during identification. In other words, one can easily commute model characteristics in vast area of operations with first principles models. Hence this approach is more popular in simulator design. In our simulator, we too have used first principles model however later multiple parameters are fine tuned with the help of practical data, as explained in Paper V [Häggblom and Ghosh, 2011]. Therefore we have used gray box modeling approach.

# 3

# System Identification

$\mathrm{T}$he aim of this section is simply to introduce a few basic concepts of parametric system identification. It neither contains the complete picture nor states any new results. More detailed explanations on these topics can be found, for example, in Söderström and Stoica [1989] and in Ljung [1999].

## 3.1   Data, model structure and selection criteria

Typically, system identification is an iterative process and is made up of three main building blocks: a) the data, b) the model structure, and c) the selection criteria. In each of these three stages personal judgment plays an important role for the final outcome.

### 3.1.1   The data $\mathcal{Z}_N$

To estimate a model we first need proper data. The quality of data depends both on the system under consideration and the experimental condition under which the data have been collected. One particular data set can be represent as a row vector

$$\mathbf{z}(t) = [y_1(t) \quad \dots \quad y_p(t) \qquad u_1(t) \quad \dots \quad u_n(t)] \in \mathbb{R}^{1 \times (p+n)} \tag{3.1.1}$$

where the system has $n$ inputs and $p$ outputs and data is collected at time $t$. A data matrix $\mathcal{Z}_N$ can construct by stacking $N$ consecutive samples, at $t = 1, \ldots, N$ on top of each other

$$\mathcal{Z}_N = \begin{bmatrix} \mathbf{z}(1) \\ \mathbf{z}(2) \\ \vdots \\ \mathbf{z}(N) \end{bmatrix} \in \mathbb{R}^{N \times (p+m)}. \tag{3.1.2}$$

It is important that the data reflect all important characteristics of the true system. The excitation signals $u(t)$ need to be chosen in such a way that the output $y(t)$ contains all the essential dynamics of the underlying system. However, often there are practical constraints (e.g. how big input amplitudes one can apply into the system, operating set-point, stability issues, etc.) which limit the experimental scope. Occasionally, when no observable input is available, the model has to be generated based on measured outputs only, i.e. $z(t) = y(t)$. This is called *time-series* modeling.

### 3.1.2  Model structure $\mathcal{M}$

The model structure selection is often considered as the single most challenging step in system identification. This problem can essentially be subdivided into three stages:

1. Selection of the **type of the model set** to consider in the modeling. In other words, this essentially means the choice between black box, grey box and physical parametric methods; between linear and nonlinear, and so forth.

2. Postulate the **size of the model set**. This consists of the choice of the possible variables and variable combinations, fixing orders and degrees of the chosen model types. By fixing these issues we can settled a model set $\mathcal{M}^*$, yet it might be quite large to handle in practice. Hence in most cases another step is required to construct a useful model with sensible approximation. Often, using prior structured information becomes handy in this stage.

3. Finally, we need to find the best possible **way to parameterize the model set** in order to estimate reasonable parameter values.

If we consider a model set $\mathcal{M}^*$ can be parameterized from a finite-dimensional parameter vector $\theta$. The model structure to which such a model belongs can be specified by the mapping, where one individual model relate to the data set $\theta$ is defined as $\mathcal{M}(\theta)$.

### 3.1.3   Selection criteria $Vn$

Almost always, in practice, there is a mismatch between the measured output and the values estimated from models. The mismatch, quantified as an error $\epsilon(t,\theta)$, is often considered as a combination of unmodeled dynamics and a measurement noise

$$\epsilon(t,\theta) = y(t) - \hat{y}(t,\theta) \qquad (3.1.3)$$

The modeling goal is to make the $\epsilon$ value as small as possible. The selection criteria we have been adopted here a scalar measure of the fit between measured and predicted values

$$V_N(\theta, Z_N) = \frac{1}{N} \sum_{t=1}^{N} l(y(t) - \hat{y}(t,\theta)) \qquad (3.1.4)$$

where $l(.)$ is a positive scalar-valued function, typically chosen as quadratic: $l(.) = (.)^2$.
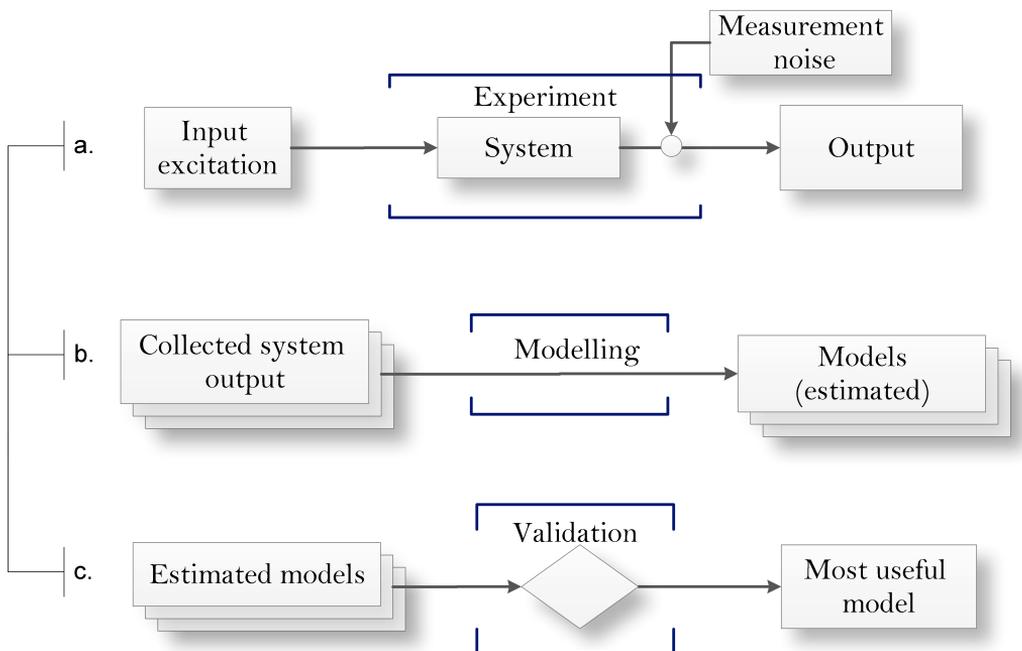


**Figure 3.1:** Typical work–flow in a system identification process
(a. experiment, b. modeling and c. model validation)

## 3.2   Design of plant experiments for identification

Data collection from an industrial process is costly and time consuming [Katayama et al., 2006]. Moreover, the process variables need to be kept within specified safety limits in normal plant operation. To make an identified model applicable to many operating regions requires vast amounts of data. The quality of a model depends on the collected process data. Therefore, limiting the excitation also limits the information available for system identification. Thus, there is a trade-off between how much one is prepared to pay for the information and the information needed for system identification. As a result, one can obtain only a limited amount and quality of experimental data [Gevers, 2005]. Therefore, it is always a challenging task to make the identification procedure efficient enough. Tailor made input excitation using process knowledge is one way to make the identification more efficient.

# Input excitation

In this chapter we will discuss various input excitation and their design methods. The idea of plant friendliness input and why it is important in system identification process are also explained here.

## 4.1  Typical input excitation

For the purpose of identification, input signals are perturbed to excite the system. The applied input signal and the resulting output of the system are measured and a model is identified from the collected data. As perturbation one can utilize many different types of signals, but step, PRBS (pseudo random binary sequence) and multisine signals are the most preferred input signals. It has been shown that advanced dedicated signals are difficult to construct [Pintelon and Schoukens, 2012]. The usual practice is to use step or PRBS signals. However, in recent years many researchers have suggested the use of multisine signals over other signals because of their plant friendly appeal [Lee, 2006; Lee and Rivera, 2006; Pintelon and Schoukens, 2012].

In addition, the excitation signals can be split up broadly into three major categories, namely general-purpose signals, optimized test signals and advanced dedicated signals, as illustrated in Figure 4.1.

a. **General purpose signal:**  This kind of signal contains an almost flat power spectrum band within the user specified frequency band. These kinds of signals are typically used for identification of an unknown system. Examples of these kinds of signals
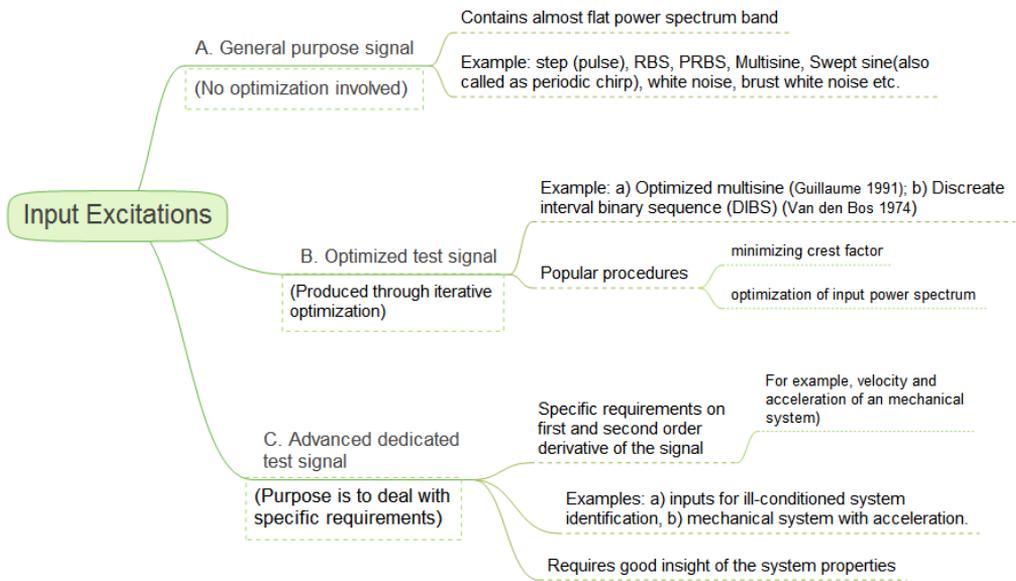
**Figure 4.1:** Input excitation types in system identification

are step, RBS, PRBS, Multisine, Swept sine (also called periodic chirp), white noise and burst white noise.

**b. Optimized input signal:** These signals are produced through optimization. In recent years, as computational power and amount of memory are not restricting factors, optimized signals have become more popular. Discrete interval binary sequence ( DIBS) [Van den Bos, 1967, 1987] and Guillaume phased optimized multisine [Guillaume et al., 1991] are examples of this kind. The optimization can involve many properties of a signal. However, *optimization of input power spectrum* and *minimizing the crest factor* of input signals are popular in the literature [Rees et al., 1992].

**c. Advanced dedicated test signal:** The aim of this kind of signal is to deal with specific system properties. The input excitation for ill-conditioned systems is a good example. For an ill-conditioned system the inputs for identification is preferred to be designed in such a way that the output is balanced. A stronger excitation is required along the relatively weak low-gain direction of the system [Koung and MacGregor, 1993]. Another example is a mechanical system with acceleration; there would be specific requirements (e.g. within specified limits) on first and second order derivative of the input signal. The input excitation needs to be designed with controlled

system properties like, velocity and acceleration [Pintelon and Schoukens, 2012]. Optimization would involve the crest factor of the first and second derivative apart from the crest factor of the input and output signal of the system. Two other popular procedures in literature are (i) simultaneous minimization of crest factor at input and output, and (ii) simultaneous minimization of peak values at the input and output. Reasonable insight of the system properties is a necessity for this signal design [Lee, 2006]. More inside information on this topic can be found in Schoukens et al. [2012].

## 4.2 Basic input excitation

Typically, any input signals are built using general purpose signals (Section 4.1.a). Here we discuss further some popular ones.

### 4.2.1 Step signal

A step signal is defined by

$$u(t) = \begin{cases} 0 & t < 0 \\ u_0 & t \geq 0 \end{cases} \tag{4.2.1}$$

where $u_0$ is the signal amplitude [Ljung, 1999; Söderström and Stoica, 1989]. Due to its simplicity it has been used extensively in practice. Often steps are used for pre-identification purpose before applying advanced dedicated signals to grasp the basic system information such as static gain, time constants, output characteristics, etc. Typically, this system information is considerably useful for advanced experiment design.

**Table 4.1:** Parameters in PRBS design

| No. | Parameter description | Value details | | Choice |
|---|---|---|---|---|
| 1 | Low dominating time constant | $\tau_L$ | | System specific |
| 2 | High dominating time constant | $\tau_H$ | | parameters |
| 3 | Number of inputs | $p$ | 4 | |
| 4 | Closed-loop response parameter | $\alpha$ | 2 | User defined |
| 5 | Settling-time parameter | $\beta$ | 3 | parameters |
| 6 | Switching time | $T_{sw}$ | | |
| 7 | Delay time (parameter for least-correlated inputs) | $D$ | | Design parameters |
| 8 | Number of bits in PRBS sequence of length $N_s = 2^{n_r} - 1$ | $n_r$ | | |

### 4.2.2  PRBS (and RBS) signal

A pseudorandom binary sequence (PRBS) is a periodic signal, which switches between two levels in a certain fashion within a user specified frequency band. A maximum length PRBS ($N = 2^{n_r} - 1$) is typically generated using a linear shift register (length $= n_r$). The user must determine the amplitude of the PRBS signal [Ljung, 1999; Söderström and Stoica, 1989]. The design parameters for the PRBS signal are summarized in Table 4.1.

The desired bandwidth of the PRBS signal can be specified according to

$$\omega_L = \frac{1}{\beta \tau_H} \leq \omega \leq \frac{\alpha}{\tau_L} = \omega_H \tag{4.2.2}$$

where $\tau_H$ and $\tau_L$ are the high and low dominating time constants of the system, respectively; $\alpha$ specifies the ratio between the open-loop and the desired closed loop time constants; $\beta \tau_H$ is the open loop settling time of the system [Rivera et al., 2009]. This frequency bandwidth can also be expressed as a fraction of the Nyquist frequency. In practice, often PRBS signal is preferred over RBS (random binary sequence) [Ljung, 1999]. A more detail discussion can be found at Section 13.3 in Ljung [1999] and at Chapter 5.3 in Söderström and Stoica [1989]. Design procedure of a PRBS is explained in Table 4.2.

### 4.2.3  Multisine signal

A characteristic of a sine signal is that all energy is located at a single frequency. In order to determine a frequency response, it is necessary to measure the output of a system for a range of frequencies. Since a full signal period is required to measure each frequency, the experiment time to measure all frequencies can be long and inefficient. Fortunately, long measurement time can be overcome while retaining all other advantages of a sinusoid by applying a multisine signal, which is a sum of sine waves with different frequencies [Pintelon and Schoukens, 2012]. A deterministic multisine signal is defined as

$$u(t) = \sum_{i=1}^{n_s} A_i \cos(2\pi f_i t + \phi_i) \tag{4.2.9}$$

where $\phi^T = [\phi_1 \phi_2 \dots \phi_{n_s}]$ is the phases of the multisine $u(t)$ and $n_s$ is the number of sinusoids. The phases $\phi_i$ and the amplitude $A_i$ of each frequency need to be chosen by the user. Although the change in phase of a multisine does not affect the power spectral density ( PSD) in the frequency domain, it alters the signal amplitude in the time domain.

Table 4.2: A1: Design procedure for PRBS signal

---

**Algorithm:** PRBS signal design procedure

---

1. Choose switching time $T_{sw}$ according to

$$T_{sw} \leq \frac{2.8 \, \tau_L}{\alpha} \tag{4.2.3}$$

where $\tau_L$ is the low dominating time constant of the system.

2. The shifting time $D$ is defined as

$$D = \frac{5 \, \tau_H}{T_{sw}} \tag{4.2.4}$$

3. The period length $N_s T_{sw}$ is determined from

$$N_s \geq \max\left( \frac{2\pi \, \beta \, \tau_H}{T_{sw}}, \, nD \right)$$

$$N_s = 2^{n_r} - 1 \tag{4.2.5}$$

where $n_r$ is the smallest integer that satisfies Eq. 4.2.5.
For the experiments with one input at a time

$$N_s \geq \frac{2\pi \, \beta \, \tau_H}{T_{sw}} \tag{4.2.6}$$

For simultaneous inputs

$$N_s \geq nD = \frac{5 \, n \, \tau_H}{T_{sw}} \tag{4.2.7}$$

in addition to Eq. 4.2.6.

4. The sampling time $T$ should be $\leq T_{sw}$. Ljung [1999] recommends

$$T \approx \frac{T_{sw}}{4} \tag{4.2.8}$$

In practice $T$ might be smaller because of other conditions.

---

Thus, one can utilize the relative phase between each of the frequency components in the multisine to affect the overall signal amplitude while keeping the energy distribution in desired frequencies [Pintelon and Schoukens, 2012]. The phase distribution of a multisine

in Eq. 4.2.9 can be linear or random or any specific form, such as

$$\phi_i = -\frac{\pi}{n_s} i(i-1) \tag{4.2.10}$$

A multisine signal with this phase distribution is referred to as a *Schroeder phase* multisine [Schroeder, 1970]. It has been shown that the Schroeder phase multisine has lower crest factor (typically 1.7) compared to a multisine signal, whose phase is either linear or random in nature [Pintelon and Schoukens, 2012].
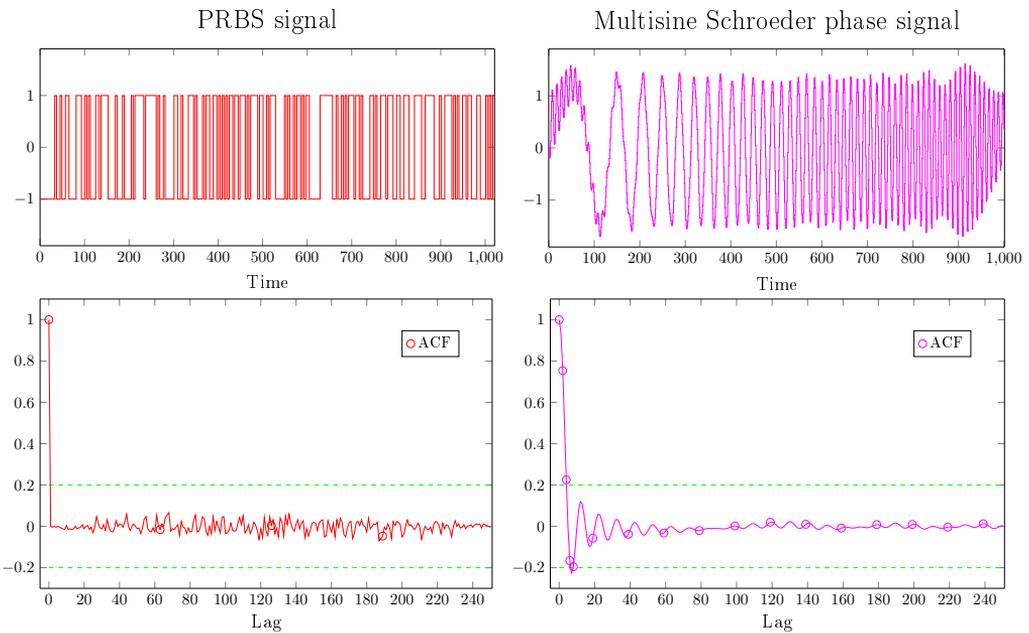


**Figure 4.2:** Auto-correlation function (ACF) of a PRBS and a multisine

The phase can also be determined by optimization to generate a plant-friendly signal. Often it is accomplished through crest factor ( CF) minimization. The relation between the phase and the crest factor of a multisine is quite complicated. As a result, the cost-function has many local minima, which makes the CF optimization problem challenging. However, using Pólya's best approximation of the Chebyshev norm it was possible to determine multisine signals with a crest factor close to the optimal value [Guillaume et al., 1991]. Another advantage of this design is to grant optimal use of the available bandwidth (measurement ranges) [Pintelon and Schoukens, 2012]. However, the crest factor is higher than that of a PRBS. Like a PRBS, multisine signals have a low degree of autocorrelation as shown in Figure 4.2.

## 4.3 Plant-friendly signals

In practice, every system has physical limitations. Such a limitation could, for example, be the sensing time interval of a sensor or the valve response time. Actuators are physical systems and have an operating limits. For example, a flow valve might catch into saturation if the control signal tries to use it beyond its maximum capacity. This might due to certain portion of the input signal remains its maximum values or the rate of change of the signal within a short period is so high that it is practically impossible accomplish. The process actuators are used in practice to introduce the perturbation into the system.

The index of plant friendliness has been expressed in many ways in control literature. A plant-friendliness of an input sequence can be defined according to the probability of change with time. The plant friendliness of an input signal is expressed as

$$\Phi = \left(1 - \frac{n_l}{N-1}\right) 100\% \tag{4.3.1}$$

where $n_l$ is the total number of switches and $N-1$ is the maximum number possible switches [Parker et al., 2001]. This means that a constant sequence is 100% plant friendly while any signal that changes at each sample is 0% plant friendly. For example, a Gaussian random sequence is a 0% plant friendly. A step input with a one switch has a plant friendliness index

$$\Phi = \left(1 - \frac{1}{N-1}\right) 100\% \approx 100\% \tag{4.3.2}$$

for large $N$ [Parker et al., 2001]. From a ttent-excitation point of view, random signals (with 0% plant-friendliness) are efficient for system identification. From a system identification point of view a constant signal (with 100% plant-friendliness) is least useful. Therefore, there is a design trade-off between the possibility to change the input signal at each point and a plant friendly input signal.

In a model based control the quality of a model is a critical factor. An inaccurate model can significantly impact on the controller performance,and finally deteriorate the quality of the plant product. It was found that about 66%–80% of the advanced control systems are not able to achieve the desired performance [Hugo, 2001], where inadequate tuning would be the main cause. Some authors have also proposed improved minimum entropy control strategies [Meng et al., 2013; Ren et al., 2014].

The definition of plant-friendliness in the above form might not suitable in formulation of optimization problem, where the equivalent form

$$\Phi_i = 1 - \frac{\sum (u_k - u_{k-1})^2}{(N-1)\max (u_k - u_{k-1})^2} \tag{4.3.3}$$

can be utilized. Here $\Phi_i \approx 1$, if $u_k \approx u_{k-1}, \forall k < N$, where $u_k, k = 1,\ldots,N$, is an input sequence [Rivera and Lee, 2003]. Plant-friendly input design is inherently multi-objective in nature [Narasimhan and Rengaswamy, 2004]. Commonly inputs are designed based on an initial estimate and after which the parameters are fine-tuned iteratively for better performance [Antoulas and Anderson, 1999].

### 4.3.1 Plant-friendliness parameters

The plant-friendliness of a signal can be quantified in many other ways. In the following we consider a few of them. The parameters are application independent measures of perturbation signal performance.

### 4.3.2 Peak factor

The plant friendliness is often expressed in terms of the peak factor ( PF) [Godfrey, 1999; Narasimhan and Rengaswamy, 2004; Van der Ouderaa et al., 1988], defined as

$$x_{PF} = \frac{x_{max} - x_{min}}{2\sqrt{2}x_{RMS}} \tag{4.3.4}$$

where $x_{max}$ , $x_{min}$ and $x_{RMS}$ denotes the maximum, minimum and root mean square ( RMS) value of the signal, respectively. The RMS of a sequence $x(k)$ is defined as

$$x_{RMS} = \sqrt{\frac{1}{N} \sum_{k=1}^{N} |x(k)|^2} \tag{4.3.5}$$

We do not consider the peak factor further.

### 4.3.3 Crest Factor

The crest factor (CF) is defined by the ratio [Barker and Godfrey, 1999; Godfrey, 1999; Godfrey et al., 2005; Guillaume et al., 2003]

$$x_{CF} = \frac{\max(|x(k)|)}{x_{RMS}} \tag{4.3.6}$$

Minimizing the peak value of the signal effectively minimizes the crest factor of the signal [Guillaume et al., 1991].

### 4.3.4 Performance index for perturbation signal (PIPS)

The PIPS for a signal is defined as

$$x_{PIPS} = \frac{\sqrt{x_{rms}^2 - x_{mean}^2}}{x_{max} - x_{min}} \ 200\% \tag{4.3.7}$$

where $x_{mean}$ is the mean value of the signal [Godfrey et al., 2005]. Both the crest factor and PIPS quantifies the goodness of energy distribution along the span of the signal. For any signal, CF can vary in between 1 and $\infty$ , whereas PIPS varies between 0 and 100%. A practical advantage of PIPS over CF is that the former always provides a finite value within 0 and 100. In this work we use both the crest factor and the PIPS as plant-friendly measures for comparing the performance of the input signals.

The crest factor of a multisine signal can be minimized by properly setting the individual phases. There are mainly two algorithms involved in this process.
(a) Clipping algorithm: the signal is generated and the pick values (higher amplitudes) are clipped in time domain. The iteration process goes on until the desired crest factor is achieved.

(b) Guillaume phase: We have considered this algorithm in our study, detail explanations can be found in next section. More details on this topic can also be found in Guillaume et al. [1991] and Schoukens et al. [2012].
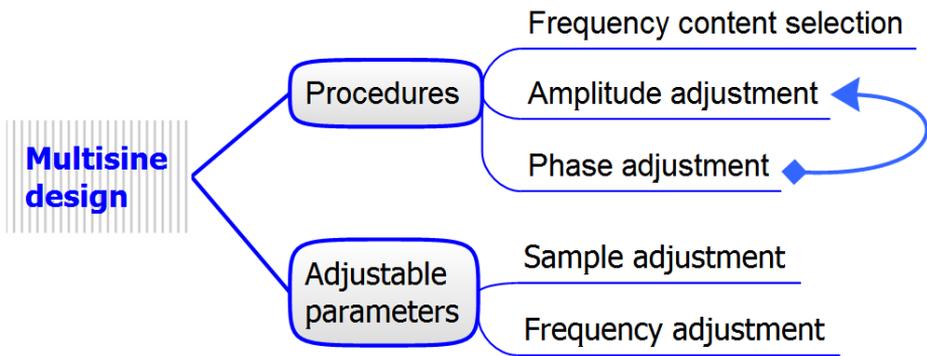


**Figure 4.3:** Designing a multisine signal

## 4.4 Designing a multisine signal

A multisine signal is a sum of harmonically related basic sine signals. A measure of input signals quality to extract the characteristics of an unknown plant is the order of persistent

excitation.

**Table 4.3:** The design parameters of different input signals [Steenis, 2009]

| Signal | Design parameters | Power spectrum** | Frequency spectrum |
|---|---|---|---|
| Step | a. Signal amplitude ($a$)<br>b. Step duration ($\Delta T$) | $\dfrac{a^2}{\sqrt{2(1-\cos\omega)}}$ | |
| Pulse | a. Pulse amplitude (1)<br>b. Pulse duration ($\Delta T$) | $\psi_p$ | $0 \leq \omega \leq \frac{2.652}{\Delta T}$ |
| Double Pulse | a. Pulse amplitude (1)<br>b. Pulse duration ($\Delta T$) | $\psi_{pp}$ | $\frac{0.8458}{\Delta T} \leq \omega \leq \frac{2.8}{T_{sw}}$ |
| PRBS | a. No. of registers ($n_r$)<br>$\quad$ for $N = 2^{n_r} - 1$<br>b. Switching time ($T_{sw}$) | $a^2 T_{sw} \frac{(N+1)}{N} \left[ \frac{\sin(0.5\omega T_{sw})}{(0.5\omega T_{sw})} \right]^2$ | $\frac{2\pi}{N T_{sw}} \leq \omega \leq \frac{2.8}{T_{sw}}$ |
| DIRBS* | a. Amplitude ($a$)<br>b. Switching time ($T_{sw}$)<br>c. Switching probability ($p$) | $a^2 T_{sw} \left[ \frac{\sin(0.5\omega T_{sw})}{(0.5\omega T_{sw})} \right]^2 ; p = 0.5$<br>(asymptotic expression) | $0 \leq \omega \leq \frac{2.652}{T}$ |
| Standard Multisine | a. Number of sinusoids ($N_s$)<br>b. Samples per cycle ($n_s$)<br>c. Sinusoid amplitude ($\alpha_i$) | $0.5 N_s (\lambda \alpha_i)^2$ | $\frac{2\pi}{N_s T} \leq \omega \leq \frac{2\pi n_s}{N_s T} \leq \frac{\pi}{T}$ |

\* Discrete Interval Random Binary Signal

\*\* Power spectrum density (psd) [†]

---

[†]The power spectral density represents the distribution of the power of the signal over the frequency interval $(-\infty, \infty)$, i.e. over both positive and negative frequencies. The power of the signal in the frequency band $(-\omega, \omega)$ is given by $\psi_\omega = \int_{-\omega}^{\omega} R_y(f)\, df$

I. A rectangular pulse of duration $\Delta T$ is defined as $p(t) = \begin{cases} 1, & 0 < t < \Delta T \\ 0, & elsewhere. \end{cases}$

The energy spectrum of the signal is $|P(f)|^2 = \left( \dfrac{sin(\pi f \Delta T)}{\pi f} \right)^2 = \Delta T^2 sinc^2(f \Delta T)$

The power spectral density $\psi_{pp}$ has maximum $\Delta T \sigma_A^2$ for $f = 0$ and zeros for $f = n/\Delta T$, where $\sigma_A^2$ is the signal variance and $n$ is a non-zero integer.

II. A double pulse of $\Delta T/2$ is defined as $p(t) = \begin{cases} 1, & 0 < t < \Delta T/2 \\ 0, & elsewhere. \end{cases}$

The energy spectrum is $|P(f)|^2 = \left( \dfrac{sin(\pi f \Delta T/2)}{\pi f} \right)^2 = \dfrac{\Delta T^2}{4} sinc^2(f \Delta T/2)$.

The power spectral density $\psi_{pp}$ has maximum $\Delta T \sigma_A^2/4$ for $f = 0$ and zeros for $f = 2n/\Delta T$. [Proakis and Manolakis, 1996]

### 4.4.1 Multisine design parameters

A multisine signal can be generated in two ways.

1. Frequency adjustment: In this process the frequencies of discrete Fourier transform ( DFT) grid are adjusted while the sample number and the clock frequency are kept constant.

2. Sample adjustment: The DFT grid frequencies are kept constants. The multisine signals are generated through adjustments in the number of samples and clock frequency. Extra attention is required to make sure that the generated signal is one period or an integer number of periods.

The design parameters of different input signals are summarized in Table 4.3 and 4.4 [Lee, 2006; Pintelon and Schoukens, 2012; Schoukens et al., 2012]. The frequency content, amplitude and phases of a sinusoidal are briefly described below. Other practical aspects in multisine design are discussed in [Schoukens et al., 2012].

**Table 4.4:** Design parameters for specific multisines in addition to (a–c) from Table 4.3 [Pintelon and Schoukens, 2012; Schoukens et al., 2012]

| Multisine type | Design parameters | In our study | Reference |
|---|---|---|---|
| Schroeder phase [Eq. 4.2.10] | d. Phase ($\phi$) as in [Eq. 4.2.10] | Exp. 9–12 | Schroeder [1970] |
| Zippered with Guillaume phase [Eq. 4.4.18] | d–e. Phase ($\phi$) and amplitude are both chosen by CF opti-mizer | Exp. 13–16 | Guillaume et al. [1991], Lee [2006] |
| Modified zippered with Guillaume phase [Eq. 4.4.21] | d–e. Phase ($\phi$) and amplitude are both chosen by CF opti-mizer f. Correlated compo-nents added in weak-gain direction | Exp. 17 | Guillaume et al. [1991], Lee [2006] |

**(a) Frequency content:** In an identification the frequency range of interest is the fre-quencies where unknown (or little known) characteristics of a system are underlying.

Input spectrum can be chosen intelligently to reduce variance errors in some frequency region of interest.  This frequency range of interest can be established through the low and high dominating time constants ($\tau_L$ and $\tau_H$ ) as

$$\omega_L = \frac{1}{\beta \tau_H} \leq \omega \leq \frac{\alpha}{\tau_L} = \omega_H \leq \frac{\pi}{T} \qquad (4.4.1)$$

where $\beta_s$ is tuned to capture settling time of the open-loop dynamics of the system and $\alpha_s$ is tuned to ensure the high frequency content of the designed signal to capture the preferred speed of closed-loop system response. For a periodic signal the lower frequency is limited to the number of samples of the signal

$$\omega_L = \frac{2\pi}{N_s T} \qquad (4.4.2)$$

Once the desired frequency range is chosen, the frequency distribution between the low and high frequency can be assigned by many methods, for example a Chebyshev, logarithmic, or a geometric sequence. Detailed description of frequency contents and its effect can be found in the literature [Pintelon and Schoukens, 2012; Rivera et al., 2009] .

**(b) Sinusoidal amplitudes:** Collected data in system identification are tainted by measurement noise. Typically, the noise effect is measured by the signal to noise ratio (SNR) of the output signal. Higher SNR makes better noise effect elimination. By applying high enough amplitude in the input of the system one can get over the noise effect, but due to physical limitations we can only have limited amplitude in the system input. Often this problem can be reduced by using long input signals. Nevertheless there is a trade-off involved between the amplitude and the duration of a signal. Amplitude limitation can further assist by avoiding signal leakage in a periodic signal, for example, using a full period length signal or $k$ periods as input; where $k$ is an integer [Pintelon and Schoukens, 2012].

**(c) Sinusoidal phases:** The phases of a multisine signal affects the amplitude in time domain, which in turn affect the plant-friendliness property of the signal.  For example a linear phase multisine, produces a signal with an impulse centered in $\tau$. The phase equation is given by

$$\phi_L(i) = -\tau \omega_i \qquad (4.4.3)$$

Both linear phase and random phase multisine generate poor plant friendly signal as shown in Figures 4.4 and 4.5. In a random phase multisine, phases are uniformly

distributed in $[0, 2\pi]$. Many design methods for generating lower crest factor multisine signals can be found in literature [Pintelon and Schoukens, 2012; Schoukens et al., 2012]. Here we discus two popular methods presented by Schroeder [1970] and Guillaume et al. [1991]. Due to complicated relation between the phase and the crest factor of a multisine signal, constructing a multisine with an optimum crest factor is not easy [Guillaume et al., 1991].
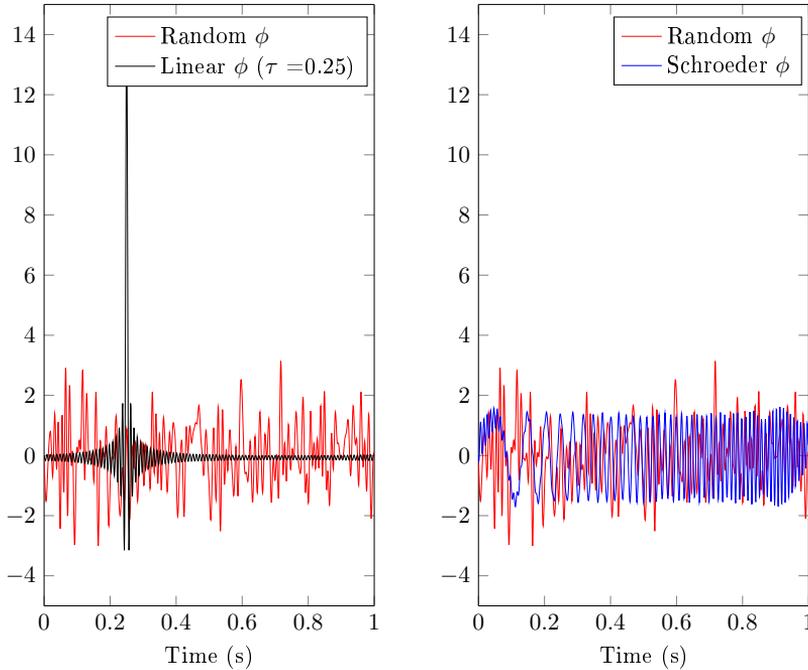


**Figure 4.4:** Multisine excitation in time domain for linear ($\phi_L(i) = -\tau\omega_i$, with $\tau = 0.25$), random (uniformly distributed in $[0, 2\pi]$) and Schroeder phases ($\phi$) [Schoukens et al., 2012]

In the Schroeder phase design, it is assumed that the multisine signal is a sum of harmonically related sine waves as

$$u(k) = \sum_{j=1}^{N} \sqrt{\left(\frac{\vartheta_j}{2}\right)} \cos\left(\frac{2\pi j}{T} k + \phi_j\right) \tag{4.4.4}$$

where $\sum_{j=1}^{N} \vartheta_j = 1$ . The Schroeder phase in closed-form formulation is given by Schroeder [1970]

$$\phi_n = \phi_1 - 2\pi \sum_{j=1}^{n-1} \vartheta_j (n-j) \qquad n = 1, 2, \cdots, N \tag{4.4.5}$$
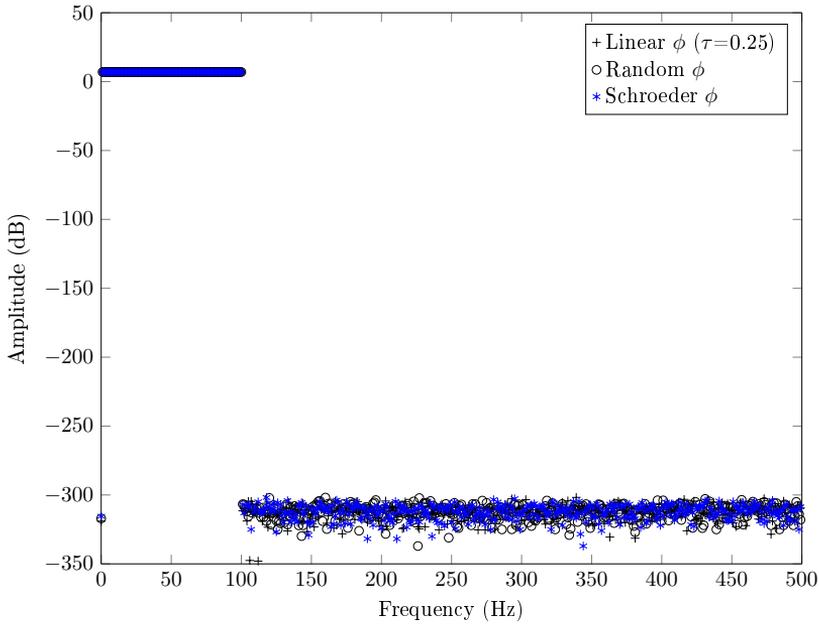
**Figure 4.5:** Amplitude diagram of a multisine excitation for linear ($\phi_L(i) = -\tau\omega_i$, with $\tau = 0.25$), random (uniformly distributed in $[0, 2\pi]$) and Schroeder phases ($\phi$) [Schoukens et al., 2012]

It is found that in a special case where the amplitudes of the multisine are all equal $\vartheta_j = 1/N$, the closed-loop phase equation Eq. 4.4.5 becomes

$$\phi_n = \phi_1 - 2\pi \sum_{j=1}^{n-1} \frac{1}{N}(n-j) \tag{4.4.6}$$

$$= \phi_1 - \pi \frac{n(n-1)}{N} \tag{4.4.7}$$

The initial phase $\phi_1$ can be any arbitrary value including zero. Now, if we set it to zero then the phase in Eq. 4.4.7 turns into the simple form

$$\phi_n = -\pi \frac{n(n-1)}{N} \tag{4.4.8}$$

This equation is known as the Schroeder phase equation.

### 4.4.2   Multisine design for a multivariable system

In a multivariable system with $n$ inputs, the dominating time constants can be defined as

$$\begin{aligned} \tau_H &= \max(\tau_1, \tau_2, \ldots, \tau_n) \\ \tau_L &= \min(\tau_1, \tau_2, \ldots, \tau_n) \end{aligned} \tag{4.4.9}$$

A priori knowledge of the system dominating time constants ($\tau_L$ and $\tau_H$) are essential for designing the primary frequency bandwidth of the desired input signal, which is defined as

$$\frac{1}{\beta \tau_H} = \omega_L \leq \omega \leq \omega_H = \frac{\alpha}{\tau_L} \tag{4.4.10}$$

In a multivariable system with $n$ inputs a multisine input $u_j(k)$ for the $j$−th channel can be defined as

$$u_j(k) = \sum_{i=1}^{n\delta} \hat{\delta}_{ji} \ \cos(\omega_i kT + \phi_{ji}^\delta) + \lambda_j \sum_{i=n\delta+1}^{n(\delta+n_s)} \sqrt{\left(2\alpha_{ji}\right)} \cos(\omega_i kT + \phi_{ji})$$
$$+ \sum_{i=n(\delta+n_s)+1}^{n(\delta+n_s+n_a)} \hat{a}_{ji} \cos(\omega_i kT + \phi_{ji}^a) \qquad j = 1,\ldots,n \tag{4.4.11}$$

where $T$ is a sampling time, $n$ is the number of channels. The sequence length $N_S = 2n(\delta + n_s + n_a)$ constitutes the Fourier coefficient, in Eq. 4.4.11. This is also a design parameter. The parameters $\delta$, $n_s$, $n_a$ are the number of sinusoids per channel and $\phi_{ji}$, $\phi_{ji}^\delta$, $\phi_{ji}^a$ are the phase angles, respectively. The frequency grid is given by $\omega_i = 2\pi i/N_s T$. The *snow effect* Fourier coefficients in higher and lower frequencies, $\hat{\delta}_{ji}$ and $\hat{a}_{ji}$ effectively reduce the crest factor of a signal [Lee, 2006].

The primary frequency bandwidth of the above multisine (in Eq. 4.4.11) is

$$\omega_L \leq \omega \leq \omega_H \tag{4.4.12}$$

where

$$\frac{2\pi n(1+\delta)}{N_s T} \leq \omega_L$$
$$\omega_H \leq \frac{2\pi n(n_s + \delta)}{N_s T} \leq \frac{\pi}{T} \tag{4.4.13}$$

Here, user has to specify $\alpha$, $\beta$, $\tau_L$, $\tau_H$, $n$ and $\delta$. At first, we need to decide the required primary bandwidth, i.e $\omega_L$ and $\omega_H$. Through the following inequality equations (4.4.14–4.4.16), one can determine the required parameters in a multisine design, such as number of sinusoids $n_s$, sampling time $T$ and sequence length $N_s$

$$(1+\delta)\frac{\omega_H}{\omega_L} \leq (n_s + \delta) \leq \frac{N_s}{2n} \tag{4.4.14}$$

$$T \leq \min\left(\frac{\pi}{\omega_H}, \frac{\pi}{\omega_H - \omega_L}\frac{n_s - 1}{(n_s + \delta)}\right) \tag{4.4.15}$$

$$\max\left(2n(n_s + \delta), \frac{2\pi n(1+\delta)}{\omega_L T}\right) \leq N_s \leq \frac{2\pi m(n_s + \delta)}{\omega_H T} \tag{4.4.16}$$

The number of sinusoidal $n_s$ in the multisine with a notch frequency $\delta$ also need to satisfy the following condition

$$(\omega_H - \omega_L) \frac{N_s T}{2\pi n} + 1 + \delta \leq n_s + \delta \leq \frac{N_s}{2n} \tag{4.4.17}$$

Few iteration steps might require for finding out appropriate design values [Lee, 2006]. Figure 4.6 shows the differences in amplitude diagram at various frequencies of an optimization algorithm based multisine and a Schroeder phase multisine signal.
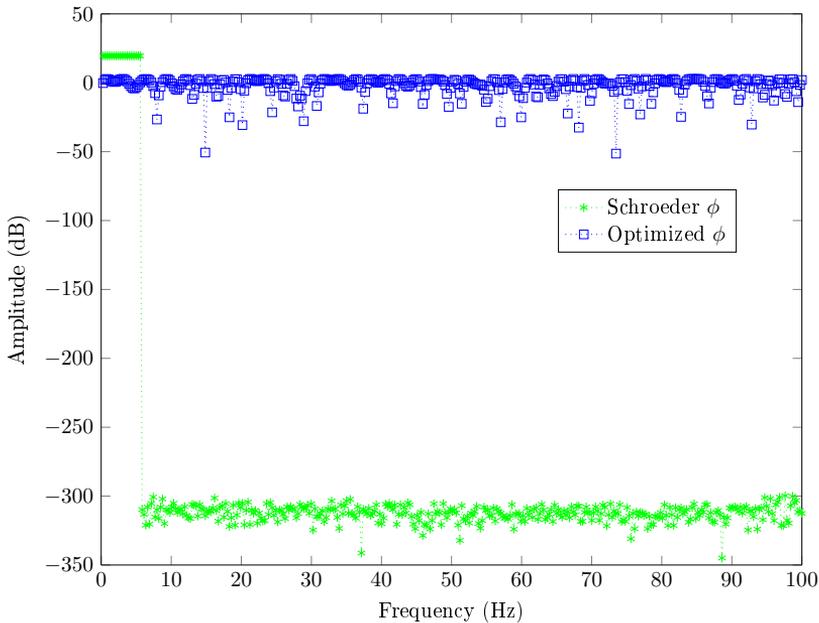


**Figure 4.6:** Amplitude spectrum of a multisine with different phases ($\phi$) [Schoukens et al., 2012]

**Design modifications for mutlisine signal**

Correlated harmonics in low gain direction have been added in this design for boosting the weak gain-direction. The idea of correlated inputs in the low-gain direction in time-domain [Zhu and Stec, 2006] has been adopted in frequency-domain using correlated harmonics. This is referred to as a *modified zippered spectrum* [Lee, 2006]. It has been shown that by applying high amplitude *correlated inputs* one can extract weaker gain information of a strongly directional system along the low-gain direction in open-loop identification [Zhu and Stec, 2006]. This type of design with correlated inputs can work for a $2 \times 2$ system, but for a MIMO system with more than two outputs the design may be inadequate. The design would be valid only for those systems where all other singular

values are much bigger than that of the weakest gain direction [Ghosh, 2014]. A detailed discussion on this topic can be found in Paper III. The Fourier coefficients $\alpha_{ji}$ for $j-$input $u_j(k)$ in Eq. 4.4.11 can be determined as

$$\alpha_{ji} = \begin{cases} \neq 0, & i = n\delta + j, n(\delta+1)+j, \ldots, n(\delta+n_s-1)+j \\ = 0, & \text{for all other i from } n\delta \text{ to } n(\delta+n_s) \end{cases} \qquad (4.4.18)$$

The *snow effect* coefficients, $\hat{\delta}_{ji}$ and $\hat{a}_{ji}$ and $\hat{a}_{ji}$, [Lee, 2006] are defined as

$$\hat{\delta}_{ji} = \begin{cases} \neq 0 & i = j, n+j, \ldots, n(\alpha+1)+j \\ = 0 & \text{for all other } i \text{ from 1 to } n\delta \end{cases} \qquad (4.4.19)$$

and

$$\hat{a}_{ji} = \begin{cases} \neq 0, & i = n(\delta+n_s)+j, n(\delta+n_s+1)+j, \ldots, n(\delta+n_s+n_a-1)+j \\ = 0, & \text{for all other i from } n(\delta+n_s) \text{ to } n(\delta+n_s+n_a) \end{cases} \qquad (4.4.20)$$

Other multisine design parameters in Eq. 4.4.11 remain same as Eq. 4.4.19 and (4.4.20) in the zippered spectrum design along with the inequalities in Eq. 4.4.16 - (4.4.17). The Fourier coefficients $\alpha_{ji}$ in the modified zippered spectrum are defined as

$$\alpha_{ji} = \begin{cases} \neq 0, & i = (n+1)\delta + j, (n+1)(\delta+1)+j, \ldots, (n+1)(\delta+n'_s-1)+j \text{ [uncorrelated]} \\ \neq 0, & i = (n+1)(\delta+1), (n+1)(\delta+2), \ldots, (n+1)(\delta+n'_s) \quad \text{[correlated]} \\ = 0, & \text{for all other } i \text{ up to } (n+1)(\delta+n_s) \quad \text{[rest]} \end{cases}$$
$$(4.4.21)$$

The Signal Spectrum of above multisine can be represented as

$$\Phi_u(j) = \begin{cases} \vartheta_L \frac{\lambda^2}{2} N_s, & j = 1, \ldots, \delta \\ \frac{\lambda^2}{2} N_s, & j = \delta+1, \ldots, n_s \\ \vartheta_H \frac{\lambda^2}{2} N_s, & j = n_s+1, \ldots, N_s \end{cases} \qquad (4.4.22)$$

while in time domain

$$u_j(k) = \sum_{i=1}^{n\delta} \hat{\delta}_{ji} \cos(\omega_i kT + \phi_{ji}^\delta) + \lambda_j \sum_{i=n\delta+1}^{n(\delta+n_s)} \sqrt{(2\alpha_{ji})} \cos(\omega_i kT + \phi_{ji})$$
$$+ \sum_{i=n(\delta+n_s)+1}^{n(\delta+n_s+n_a)} \hat{a}_{ji} \cos(\omega_i kT + \phi_{ji}^a) \qquad j = 1, \ldots, n$$
$$(4.4.23)$$

The zippered (modified) multisine signals are generated by using the CRIDENT toolbox [Lee, 2006; Lee and Rivera, 2006]. As we are comparing various input excitation we have tried to keep the experiment length relatively equal.

### 4.4.3   Summary of multisine design

As for PRBS design, we need to have some a priori knowledge of the system for multisine design. A frequency spectrum of interest can be determine by the prior system knowledge $\alpha$, $\beta$ and dominating time constants $(\tau_L, \tau_H)$ of the system using the relation

$$\frac{1}{\beta \tau_H} = \omega_L \leq \omega \leq \omega_H = \frac{\alpha}{\tau_L} \tag{4.4.24}$$

Moreover, for an $n$ inputs system, all the frequencies have to fit inside the bandwidth of the underlying system, i.e.,

$$\omega_H - \omega_L \leq n \frac{2\pi}{N T_s}(n_s - 1) \tag{4.4.25}$$

The design practices have been discussed in Paper III and IV [Ghosh et al., 2016; Häggblom and Ghosh, 2015] and guidelines can also be found in Rivera et al. [2007].

**Standard multisine design**

Two important parameters in multisine design are the period length $N_s T$ and the number of frequencies ($n_s$). The normal (non-zippered) design algorithm is summarized in Table 4.5.

Table 4.5: **A2:** Design procedure for standard (non-zippered) multisine signal

---

**Algorithm:** Standard multisine signal design procedure

---

1. Because of alias at higher frequencies than the Nyquist frequencies, the sampling time $T$ must satisfy

$$T < \frac{\pi}{\omega_H} \quad \text{(usually } \ll \text{)} \tag{4.4.26}$$

   Usually $T$ is decided by other factors than *signal design*.

2. Choose smallest integer $N_s$ according to

$$N_s \geq \frac{2\pi}{T \omega_L} \tag{4.4.27}$$

3. For equally spaced frequencies choose smallest integer $n_s$ such that

$$n_s \geq \frac{N_s T}{2\pi} \omega_H \tag{4.4.28}$$

---

**Zippered multisine design**

The zippered spectrum with Guillaume phase [Guillaume et al., 1991] is designed by exciting only one input channel at one frequency grid, so that each input channels maintain orthogonality and thereby lower correlated inputs. Furthermore, in order to get proper direction balance in output, the low-gain direction were emphasized in zippered frequencies, which is referred as 'modified zippered' signal [Lee, 2006]. The zippered multisine design with $n$ signals and notch frequency $\delta$ (= integer, is assumed to be known) is summarized in Table 4.6.

**Table 4.6: A3:** Design procedure for multisine zippered signals

**Algorithm:** Zippered multisine signal design procedure

1. Choose sampling time $T$ according to

$$T < \frac{\pi}{\omega_H} \quad \text{(usually} \ll \text{)} \tag{4.4.29}$$

   where $\omega_H$ is the highest frequency of interest.

2. Choose smallest integer $N_s$ according to

$$N_s \geq \frac{2\pi}{T\omega_L} n(1 + \delta) \tag{4.4.30}$$

3. Choose smallest integer $n_s$ that satisfies

$$n_s \geq \frac{N_s T}{2\pi n} \omega_H - \delta \tag{4.4.31}$$

Note that the zippered design for MIMO systems gives very long period time period $N_s T$ because of the factor $n$ in Eq. 4.4.30. This also means that the experiment time would be long. In case of multisine design $T \leq 0.6/\omega_H$ might also be a good choice (as we have found in case of PRBS) .

# Identification of ill-conditioned systems

In this chapter we are going to discuss directionality and the identification of ill-conditioned systems.

## 5.1 Introduction

Identifying a MIMO system is a challenging task. Often this problem is tackled by considering a set of MISO (multi-input single-output) schemes, one for each output signal (i.e. a sub-transfer function) instead of a full MIMO scheme. The frequency response and model uncertainty regions are then examined for each MISO cases separately [Hakvoort, 1994; Van den Hof et al., 1995].

In geometrical terms, directionality means that the length of the output vector strongly depends on the direction of the input vector. Now let us consider an example to present the effects of directionality in a system.

In a distillation column, due to the strong directionality present in the system, a normal excitation with input changes in same direction, i.e. $\Delta L/\Delta V > 0$ ( e.g., $[1\ 1]^T$) produces little change in output concentrations ($y_D,\ x_B$). On the other hand, the same amplitude change in opposite directions, i.e. $\Delta L/\Delta V < 0$ (e.g $[1\ -1]^T$), produces significant changes in outputs. These scenarios are shown in Figure 5.1.
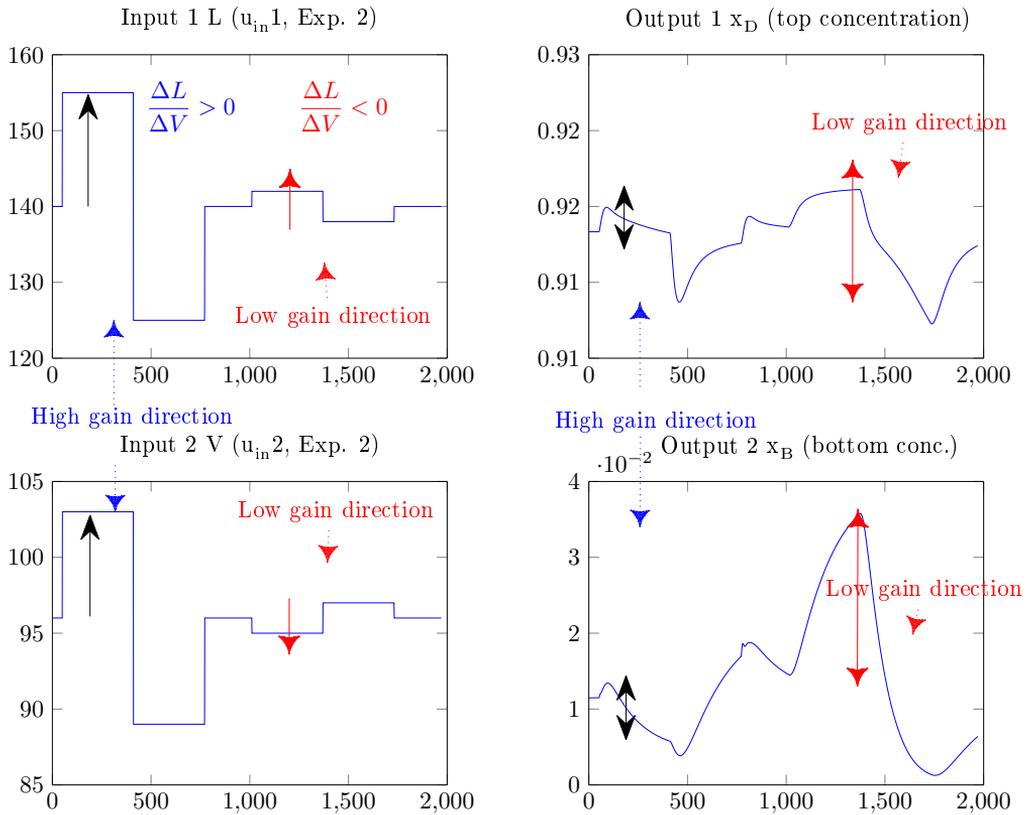
**Figure 5.1:** Low- and high-gain directions in a distillation column: inputs ( L , V , left) and outputs ($x_D$ , $x_B$ , right )

## 5.2   Directional inputs

To construct a control-relevant model, the system has to be excited adequately in all gain directions. It is especially important to excite in weak gain directions, to limit the relative uncertainties associated with these gain directions [Häggblom and Böling, 1998].  In practice it is difficult to extract system information in the low gain direction. If one input at a time is perturbed, or if all inputs are perturbed simultaneously but independently of each other, the system is usually not excited along the gain directions. In both cases, the weak gain directions tend to excite insufficiently. One can solve this problem in two different ways: a) apply especially designed inputs based on the gain directions of the system in open-loop identification, and b) apply set point changes in different directions in closed-loop identification. In both cases, some prior information of the system is needed. Excitations can be introduced one direction at a time or all directions simultaneously.

Consider a singular value decomposition ( SVD) of the steady-state gain matrix

$$G(0) = W\Sigma V^T \tag{5.2.1}$$

The input $u^i = v_i \sigma_i^{-1}$ will now produce the output $y^i = w_i$ , with $|y^i| = 1$. Where $v_i$ and $w_i$ are the $i$−th column vector of $V$ and $W$ matrix, respectively, in Eq. 5.2.1. To properly excite all gain directions"$i$", $i = 1, 2, \ldots, n$ , the inputs $u^i$ should vary between $u^i_-$ and $u^i_+$ , defined as

$$\begin{aligned} u^i_- &= -\sigma_i^{-1} \, v_i \\ u^i_+ &= +\sigma_i^{-1} \, v_i \end{aligned} \tag{5.2.2}$$

This can be achieved by any kind of input signal (step sequence, PRBS or sinusoidal). The common practice is to use PRBS signals for input excitation, one at a time or simultaneously while maintaining least correlation among the inputs. However, this cannot excite the system in all directions adequately, especially in the low-gain direction [Häggblom and Böling, 1998; Häggblom and Böling, 2013].

## 5.3 Ill-conditioned systems

In the case of an ill-conditioned system, the outputs with normal inputs are dominated by high gain direction information. As a result, it is almost impossible to extract low gain information using normal input excitation. Therefore, one has to design and apply inputs in such a way that the system output would also contain lower (weaker) gain informations. Such information is normally important when designing a controller based on an identified model of the true system. Here we have two other examples of a $2 \times 2$ ill-conditioned system which has been studied extensively in literature. The reader may refer to Appendix B for more examples of ill-conditioned systems.

### Example 1

This system is a distillation column [Skogestad et al., 1988]. Transfer matrix of the system is

$$G(s) = \begin{bmatrix} \frac{87.8}{194s+1} & \frac{-87.8}{194s+1} + \frac{1.4}{15s+1} \\ \\ \frac{108.2}{194s+1} & \frac{-108.2}{194s+1} - \frac{1.4}{15s+1} \end{bmatrix} \tag{5.3.1}$$

Using singular value decomposition (SVD) of the steady-state transfer function matrix, Equation (5.2.1) yields

$$\Sigma = \begin{bmatrix} 197.21 & 0 \\ 0 & 1.3914 \end{bmatrix} \quad W = \begin{bmatrix} -0.6246 & -0.7809 \\ -0.7809 & -0.6246 \end{bmatrix} \quad V = \begin{bmatrix} -0.7066 & -0.7077 \\ -0.7077 & -0.7066 \end{bmatrix} \quad (5.3.2)$$

The magnitude of the singular values $diag(\Sigma)$ are far apart. The condition number $\kappa = \frac{\sigma_1}{\sigma_2} > 140$ and the system contains a strong directionality.

## Example 2

This is another example of an $2 \times 2$ ill-conditioned system. This is an example of a heat exchanger system [Jacobsen and Skogestad, 1994]. The transfer function matrix ( $G(s)$) of the system is

$$G(s) = \frac{89.243}{(100s + 1)(2.439s + 1)} \begin{bmatrix} -21(4.76s + 1) & 20 \\ & \\ -20 & 21(4.76s + 1) \end{bmatrix} \quad (5.3.3)$$

SVD of the steady-state transfer function matrix of Eq. 5.3.3 gives

$$\Sigma = \begin{bmatrix} 3658.96 & 0 \\ 0 & 89.24 \end{bmatrix} \quad W = \begin{bmatrix} -0.7071 & -0.7071 \\ -0.7071 & 0.7071 \end{bmatrix} \quad V = \begin{bmatrix} 0.7071 & 0.7071 \\ -0.7071 & 0.7071 \end{bmatrix} \quad (5.3.4)$$

Strong directionality is present in the system. The condition number $\kappa = 40$.

## 5.4   Input design methods

The output of an ill-conditioned system with random input excitation is strongly aligned along the high-gain direction of the system and contains merely little information of the low-gain direction. In order to avoid this, some design methods have been proposed in literature, which are summarized in Table 5.1 * Here these design methods are discussed briefly.

Almost all design methods and related studies except [Koung and MacGregor, 1994] and [Li and Lee, 1996], are focused on $2 \times 2$ systems. Some design methods are not as effective for $n \times n$ systems where $n > 2$, as they are for $2 \times 2$ systems. Some process systems used in literature for these studies are summarized in Table 5.2.

---

*please refer Appendix B for model details.

### 5.4.1 Rotated inputs

Koung and MacGregor [1993] have proposed a design method using rotated inputs. The rotation is determined based on the rotation matrix $V$ from SVD of the steady-state transfer function matrix of the underlying system. The designed input vectors are aligned with the system's high– and low-gain directions and able to excite both high– and low-gain directions in a balanced way. Another design method based on rotated signals has been presented in Conner and Seborg [2004] using PRBS signals.

Subspace identification has gained popularity in recent years due to its numerical appeal, but the possibilities to use prior system knowledge is limited. Determining the correct system order of an ill-conditioned system is often a challenge. However, this problem can overcome through rotated inputs. Misra and Nikolaou [2003] have laid out a design based on a particular *rotated angel*. However no method, except *trial and error*, is given to find the angle. Obviously, this is a drawback of this design method. Moreover, Micchi and Pannocchia [2008] have argued that finding a rotation angle is a challenge in this method.

The $W$ and $V$ matrix from Equation (5.3.2) can be re-written as [Misra and Nikolaou, 2003]

$$W_\phi = \begin{bmatrix} \cos\phi & -\sin\phi \\ \sin\phi & \cos\phi \end{bmatrix} \quad V_\theta^T = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \qquad (5.4.1)$$

One can identify an ill-conditioned system order by applying a random input with other input *rotated* as $u_2 \approx u_1 \cot\theta$. The angle $\theta$ has to be found by iteration through examining the maximum separation of the pair of singular values.

It has been found that the effectiveness of the rotated angle method is very sensitive to the accuracy of the applied angle which is determined by trial and error. To avoid multiple trials to find correct angle(s), random setpoints in closed-loop test identification have been suggested by Micchi and Pannocchia [2008]. Also, some subspace based algorithms might not be suitable because of correlation between the inputs. For example the orthogonal projection method [Huang et al., 2005], is such a method. In case of $2 \times 2$ systems, finding the angle between two inputs is much easier, but for systems with more inputs it becomes complicated.

**Table 5.1:** Main characteristics of signal design methods for ill-conditioned process identification

| Design method | A-priori knowledge | Calculation principle | Signal Amplitude | Signal Spectrum | Reference | Input excitation |
|---|---|---|---|---|---|---|
| Rotated signal | Steady-state gain matrix | Rotation by $W$ of SVD($G(0)$) matrix | Four level | Fairly flexible | Conner and Seborg [2004]; Häggblom and Böling [1998]; Koung and MacGregor [1993] | Step, PRBS |
| Correlated high amplitude in low-gain method A | Sign of low-gain direction | Concatenated signal | Four level | Fairly flexible | Zhu [2001] | PRBS |
| Correlated high amplitude in low-gain method B | Sign of low-gain direction | Combined signal | Four level | Fairly flexible | Zhu [2009] | PRBS |
| SOH with modified zippered power spectrum | Steady-state gain matrix | Phase optimization and Rotation by SVD rotation matrix | Continuous | Flexible | Lee [2006]; Lee and Rivera [2006]; Rivera et al. [2009] | PRBS, multi-sine |
| Pseudo random trinary sequence (PRTS) | Sign of low-gain direction | Finite field arithmetic | Three level | Inflexible | Tan [2009] | PRTS |

### 5.4.2   Correlated signal method A

A relatively simpler design method for a $2 \times 2$ highly-interactive system has been presented in Zhu [2001]. This method has two stages: a) a normal random input signals to produce

**Table 5.2:** Main characteristics of the models originally employed to test the signal design methods for ill-conditioned process identification

| Design method | Model reference | Model order, size | RGA ($\lambda$), Cond. No. ($\kappa$) | Identified model structure | Identification reference |
|---|---|---|---|---|---|
| Rotated signal | Wood and Berry [1973] | FOPTD[†] $2 \times 2$ | 2.009, 7.48 | Low-order ARX, FIR | Conner and Seborg [2004]; Koung and MacGregor [1993] |
| Correlated high amplitude in low-gain method A | Jacobsen and Skogestad [1994] | $2^{nd}$ order $2 \times 2$ | 35.068, 141.73 | ARMAX, BJ, SS MOSEP | Zhu [2001] |
| Correlated high amplitude in low-gain method B | Jacobsen and Skogestad [1994] | $2^{nd}$ order $2 \times 2$ | 35.068, 141.73 | ARMAX, SS MOSEP | Zhu [2009] |
| SOH with modified zippered power spectrum | Jacobsen and Skogestad [1994] | $1^{st}$ order $2 \times 2$ | 35.068, 141.73 | ARX, NARX | Lee [2006]; Lee and Rivera [2006]; Rivera et al. [2009] |
| PRTS[‡] | Ogunnaike et al. [1983] | SOPTD[§] $3 \times 3$ | 2.008, 3540 | Frequency domain system identification | Tan [2009] |

∗ First Order Plus Time Delay
† Pseudo Random Trinary Sequence
‡ Second Order Plus Time Delay

high-gain rich outputs, and b) using the high-gain system information from an earlier step, a strongly–correlated high–amplitude inputs (almost identical) is applied in the low

gain direction to extract low-gain rich outputs.

The main advantages of the this method are fast and simpler identification design. This method is best suited for a $2 \times 2$ systems, especially distillation columns, as shown in Eq. 5.3.1 and 5.3.3, where typically, high-gain direction is aligned close to $[1 \quad 1]^T$ vector, while low-gain direction is aligned towards $[1 \quad -1]^T$ (i.e. changing inputs in opposite direction). In case of a distillation column operating with L-V configuration (as shown in Figure 5.1), inputs and outputs are $(L, V)$ and $(y_D, x_B)$, respectively.

### 5.4.3   Correlated signal method B

Zhu [2009] has presented an extension to the method described in Section 5.4.2 [Zhu, 2001]. Here, both uncorrelated random with low-amplitude and correlated high-amplitude signals are added and applied together.

### 5.4.4   Frequency based design

The above idea, from Sections 5.4.2 and 5.4.3, has been employed in frequency domain design with multisine signals [Lee, 2006]. Multisine signals have been designed by introducing correlated harmonics in power spectrum in order to compensate low-gain direction[Lee, 2006]. Rivera et al. [2009] have presented another method to generate plant-friendly signals while minimizing the crest-factor (CF) of the signal by optimizing both phases and amplitudes of the frequency harmonics of the designed signal. In this procedure, Schroeder phase Equation (4.2.10) was used as the initial phase in the optimization. Finally, the idea of snow effects in order to minimize the crest factor of the signal, as proposed in Guillaume et al. [1991], was used.

A similar method using frequency domain based design has been used with Wiener modeling for a non-liner system [Guillaume et al., 2003]. Another input design method based on correlated harmonics in frequency domain using *Galois field*, has been presented in [Tan, 2009]. The method used a three-level signal called PRTS (*pseudo-random ternary signal*) that can compensate the low-gain direction. However, this method has two limitations. The user does not have any control over the frequency content, unlike other design methods, and knowing the low-gain direction is also a prerequisite.

## 5.5 Evaluation of inputs

In our study, we have considered both designs by rotated signals as well as the methods based on frequency domain design. In Paper I and Paper II, we have studied rotated signal based designs and frequency domain based design, respectively. We have compared and studied all the significant design methods and summarized in Paper III. In Paper IV, we have focused on a few design methods and their design aspects in practical scenarios. To compare the effectiveness of input signals we have considered both fit percentage and projection variance of output data in gain directions.

### 5.5.1 Fit percentage

The fit percentage is considered as a measure of goodness of the estimated models from the measured data. The normalized root mean square error (NRMSE) is considered to determine the goodness of fit for each output of the multiple-input single-output (MISO) model estimated in identification. The fit percentage value is calculated as

$$F_p[\%] = \left(1 - \frac{\sqrt{\sum_i(\hat{y}_i - y_i)^2}}{\sqrt{\sum_i(y_i - \bar{y}_i)^2}}\right) 100\% \qquad (5.5.1)$$

where $y_i$ and $\hat{y}_i$ are the measured and estimated output, respectively, and $\bar{y}$ is the mean value.

### 5.5.2 Projections of output data in gain directions

The standard deviation of the projections of system output along the gain-directions illustrates how well the output is balanced in various gain directions. These projections at each sampled data points are calculated according to

$$y_i^\times(k) = w_i^T y(k) \qquad (5.5.2)$$

where $y_i^\times$ is the projection of output $Y$ along the $i$–th gain-direction and $w_i$ are the columns of the $W$ matrix from the singular value decomposition of the system gain matrix using Eq. 5.2.1.

# Case studies

Input signals can be constructed with and without information about gain directions of a system. An input signal, as explained earlier in Section 4.2, can for example be a step, a PRBS or a multisine. In this contribution, we have considered various design methods to construct efficient input signals using all three types of excitations.

## 6.1 Introduction

Typically, a designer has to decide the amplitude and the desired frequency bandwidth of a signal (e.g. PRBS or multisine). The phase of a multisine can be used to get the desired amplitude of the signal in time domain without changing the power spectral density of the signal in frequency domain. Phase can be modified by an optimization to obtain more plant-friendly multisine inputs [Lee, 2006]. Typically, Schroeder phase distribution [Schroeder, 1970] has been used as an initial choice in optimization. The relation between phase and crest factor (CF) of a signal is complicated. As a result, constructing globally optimized multisine signal with minimum crest factor (CF) is not easy, however, it is possible to get a close approximation using Guillaume phase distribution [Guillaume et al., 1991].

### 6.1.1 Input excitation design

We have used various input design methods to construct efficient input excitations. Based on signal types, as explained in Section 4.2, we can divide the experiments into three

categories, namely, (a) step based signals (Exp. 1 & 2), (b) PRBS based signals (Exp. 3–8), and (c) multisine based signals (Exp. 9–17). In each set, some of the designs are standard designs, some are based on informations about gain directions. In addition, the designed excitations can be sequential (one at a time) or simultaneous (applied together). In case of PRBS inputs, design guidelines given by Gaikwad and Rivera [1996] have been used as a starting point. The design algorithms for PRBS (Table 4.2) and multisine are used, as explained in Section 4.2. All these experiments are summarized in Table 6.1.

**Table 6.1:** Designed experiments based on input perturbations

| Exp. | Input perturbation description | Acronym | Direction information |
|------|-------------------------------|---------|----------------------|
| 1 | Sequential step changes of inputs one at a time | stepSeq | No |
| 2 | Step changes in gain directions | stepDir | Yes |
| 3 | Sequential PRBS perturbation of inputs one at a time | prbsSeq | No |
| 4 | Simultaneous uncorrelated PRBS perturbation of all inputs | prbsUnc | No |
| 5 | Sequential PRBS perturbation of gain directions | prbsSeqDir | Yes |
| 6 | Simultaneous PRBS perturbation of gain directions | prbsSimDir | Yes |
| 7 | Sequential high-amplitude correlated PRBS and low-amplitude uncorrelated PRBS | prbsDirCorrSeq | Yes* |
| 8 | Simultaneous high-amplitude correlated PRBS and low-amplitude uncorrelated PRBS | prbsDirCorrSim | Yes* |
| 9 | Sequential Schroeder phase multisine perturbation | msSrdSeq | No |
| 10 | Sequential Schroeder phase multisine perturbation of gain directions | msSrdSeqDir | Yes |
| 11 | Sequential Guillaume phase multisine perturbation | msGlmSeq | No |
| 12 | Sequential Guillaume phase multisine perturbation of gain directions | msGlmSeqDir | Yes |
| 13 | Simultaneous Schroeder phase multisine perturbation | msSrdSim | No |
| 14 | Simultaneous Schroeder phase multisine perturbation of gain directions | msSrdSimDir | Yes |
| 15 | Simultaneous Guillaume phase multisine perturbation | msGlmSim | No |
| 16 | Simultaneous Guillaume phase multisine perturbation of gain directions | msGlmSimDir | Yes |
| 17 | Guillaume phase multisine modified zippered signal | msGlmZipm | Yes* |

* only low-gain directional information are used.

The design in experiment 1 (stepSeq) is the simplest method. Here steps are applied to the inputs with both positive and negative amplitude. Although the method is simple, it provides fundamental information of the system, which is useful in advanced designs. The amplitude of the signal is a user choice. Typically it is determined based on prior knowledge, or obtained e.g. by trial and error. Signals from the first experiment have modified with system gain directions and applied in Experiment 2 (stepDir). The gain directions are estimated from the step tests in experiment 1, and applied to the inputs as explained in Section 5.2. For fair comparison of various designs, we have designed

reasonably equal experiment lengths for all experiments. Each sub-figure in Figure 6.1 shows inputs of each experiments.
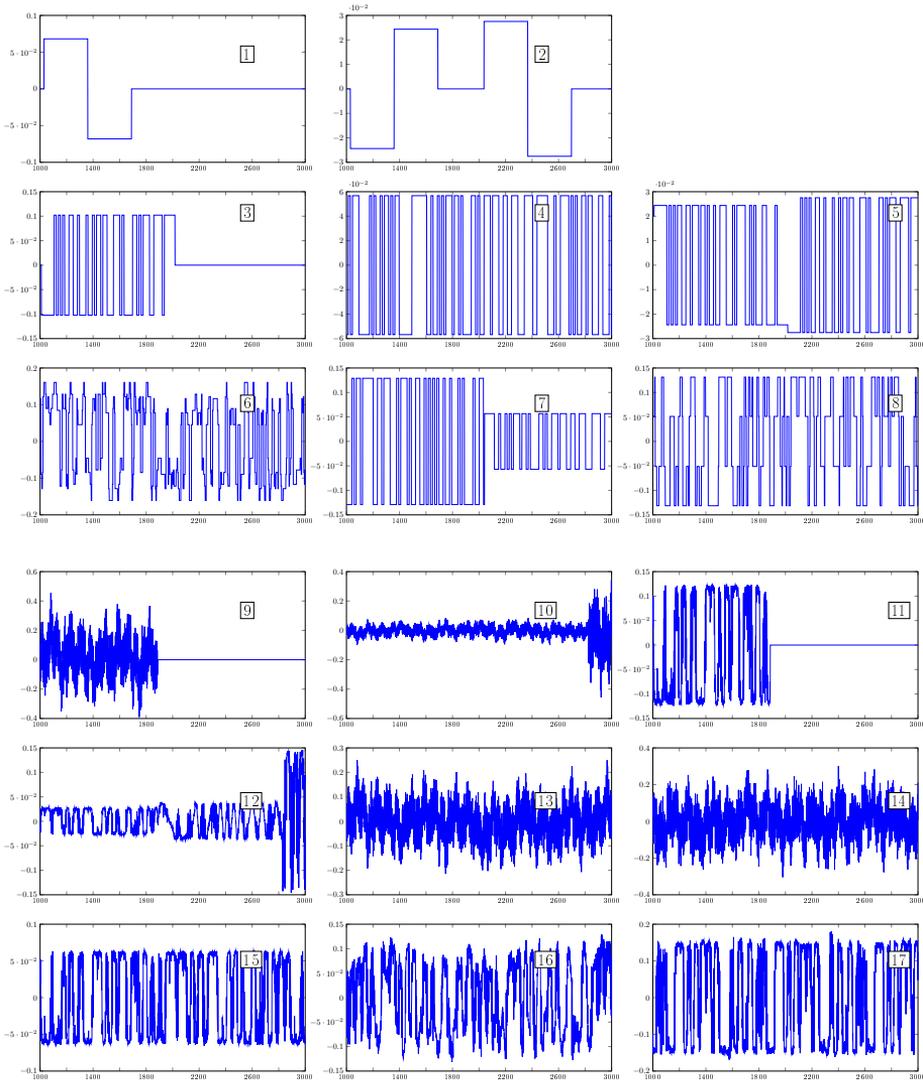


**Figure 6.1:** Input excitation signals— step, PRBS and multisine (selective part)
[Ghosh et al., 2016]

Experiment 3 (prbsSeq) has been designed using PRBS signals in each input, one at a time. The PRBS signals are designed based on the design parameters specified in Table 4.1. In experiment 4 (prbsUnc), all inputs are excited simultaneously, in an uncorrelated way. Uncorrelated signals are obtained by time shifting the PRBS signals suitably among the four inputs [Lee, 2006]. In experiment 5 (prbsSeqDir), the PRBS

signal from the third experiment is applied to the gain directions with suitably adjusted amplitudes. In experiment 6 (prbsSimDir), the signals of experiment 4 are applied to the various gain directions with suitably adjusted amplitudes.

Experiments 7 and 8 are motivated by the design method in Zhu [2009], where only the low-gain direction is emphasized with an explicit excitation combined with standard uncorrelated excitation. Thus, this method requires knowledge of low-gain direction of the system. We have used the estimated low-gain direction obtained from Exp. 1 (stepSeq). In accordance with Zhu [2009], a correlated high-amplitude signal was applied in the low-gain direction of the system while the high-gain direction was excited by uncorrelated signals of lower amplitude. In Both experiments, PRBS signal were used and the signals were scaled so that output magnitudes were similar to the ones in the other experiments. The difference between experiment 7 (prbsDirCorrSeq), and experiment 8 (prbsDirCorrSeq) is that the inputs are applied sequentially in 7 and simultaneously in 8.

In experiments 9–17 multisine inputs are used. In a few experiments 9–12, the inputs are applied sequentially, while in experiments 13–16, they are applied simultaneously, either directly to the inputs or to the gain directions. In experiment 9 (msSrdSeq), the multisine signal is designed with Schroeder phase distribution [Schroeder, 1970] and is applied directly to the inputs, one by one. In experiment 10 (msSrdSeqDir) the same signals with suitably adjusted amplitudes, are applied to the various gain directions. Experiments 11 (msGlmSeq) and 12 (msGlmSeqDir) are similar to experiment 9 and 10, but instead the Guillaume phase distribution is used. The phase is determined based on an optimization algorithm for minimal crest-factor of the signal by iterative $p-$norm optimization [Guillaume et al., 1991]. The designs in experiment 13 (msSrdSim) and 14 (msSrdSimDir) are similar to those in experiment 9 (msSrdSeq) and 10 (msSrdSeqDir). Experiment 15 (msGlmSim) and 16 (msGlmSimDir) are similar to the design in experiments 11 (msGlmSeq) and 12 (msGlmSeqDir), but all excitations are applied simultaneously instead of sequentially.

In experiment 17 (msGlmZipm), mulstisine zippered signals with Guillaume phase distribution has been used. The CRIDENT toolbox [Rivera and Lee, 2003] has been used in designing the multisine signals. This is an adaptation in frequency domain with correlated frequencies [Rivera et al., 2007], of the earlier methods (as in experiment 7 and 8) in time-domain Zhu [2009]. In this design, low-gain direction of the system is compensated, whereas algorithm generated certain low and high frequency contents are added for minimizing the crest factor of the signal [Lee, 2006].

### 6.1.2 Modeling and cross validation

MATLAB's Simulink and System Identification Toolbox have been used for modeling [Ljung, 1999, 2012]. Various model structures (e.g. transfer function matrix model, state space model, polynomial model) have been tried out. But there was no significant difference in performance. For easy comparison with the true system, we have chosen the *transfer function matrix* structure. It has been shown that the lower gain dynamics is much faster than the dynamics of the higher gain direction in distillation columns [Skogestad et al., 1988]. Hence, in order to consider both slow and fast dynamics of the system in estimated model, we have considered two poles model with time delay for each transfer function

$$G_{ij} = \frac{K_p \ e^{-sT_d}}{(1 + sT_{p_1})(1 + sT_{p_2})} \tag{6.1.1}$$

where $G_{ij}$ is the $i^{th}$ row and $j^{th}$ column element of the transfer function matrix of the system. To avoid impractical delay parameters in the numerical optimization, which was often an issue, a realistic upper limit of 20 minutes was used for the delay values.

The designed inputs are applied and the system response (i.e. outputs) are measured for each experiment. Models are determined from the experimental data and the model that fits the data most is chosen. The performance of a model estimated from an experiment is also validated through data from other experiments not used in estimating the model.

A two-pole with time delay model for each transfer function is fitted to each experiment as Equation (6.1.1). The performances of the designed inputs are then verified with each data-set, separately. As we have designed input signals using various methods, as in Table 6.1, we could consider them a good set of validation data for rest of the experiments. Hence, the merits of the input excitations are also evaluated through cross-validations with the data of the other designed experiments.

## 6.2 Experiment design example - a distillation column system

We have considered design methods as explained in Section 6.1.1 for a 2×2 distillation column system. This nonlinear simulator is designed to complement a pilot-scale distillation column system at Process Control Laboratory (Åbo Akademi University). Therefore, by tuning a number of parameters of the simulator, the behavior of the simulator is adapted to that of the real column. A schematic of a two-product distillation column is shown in 6.2.
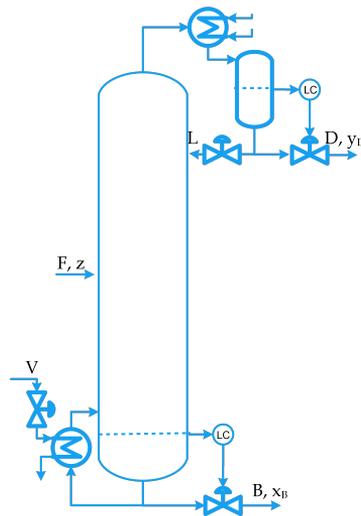
**Figure 6.2:** Schematic of distillation column system

Apart from the projections of output data on gain directions, as explained in Section 5.5.2, the spread of outputs produced by designed inputs can be characterized by the determinant of the correlation matrix of the outputs, i.e.,

$$\Upsilon = \det(\text{corr}\,[y_1 \quad y_2 \quad \ldots \quad y_n]) \tag{6.2.1}$$
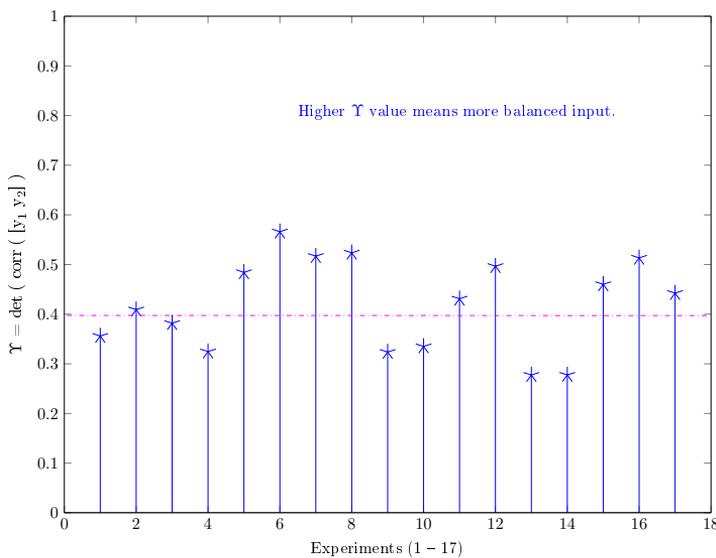


**Figure 6.3:** Calculation of determinant of output-correlation ($\Upsilon$) of designed experiments in simulator

Figure 6.3 shows determinant of output-correlation ($\Upsilon$) values for all the designed experiments. Based on the $\Upsilon$ value, we can distinguish between the experiments which produce more balanced outputs and which do not. A few screen-shots of our designed input excitations and corresponding outputs are shown in Figures 6.4 to 6.6.
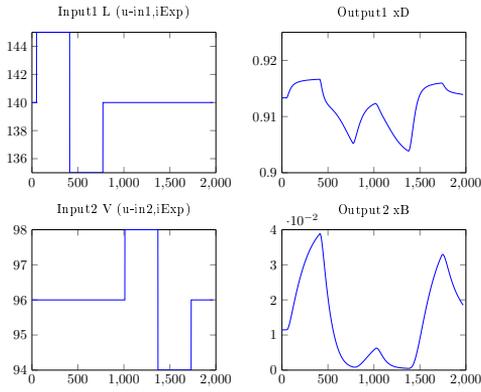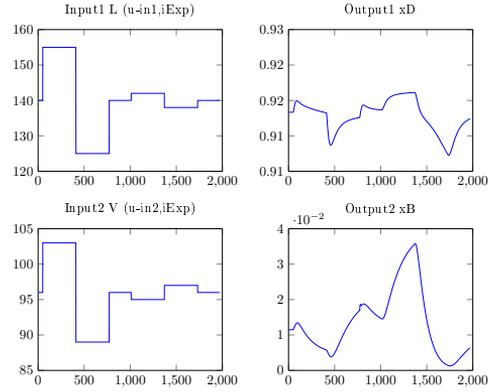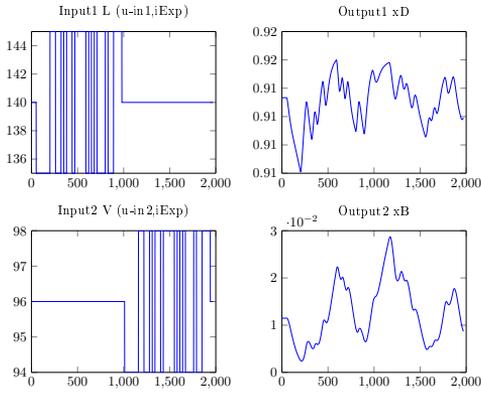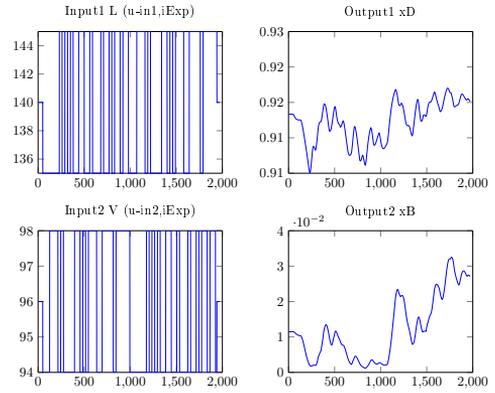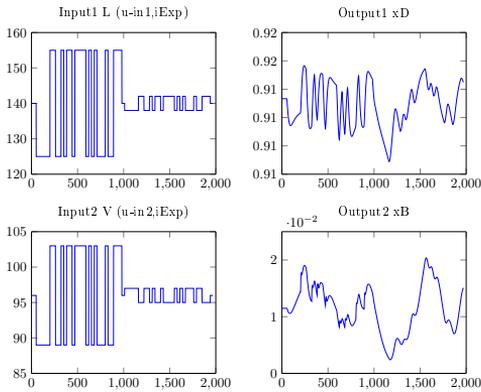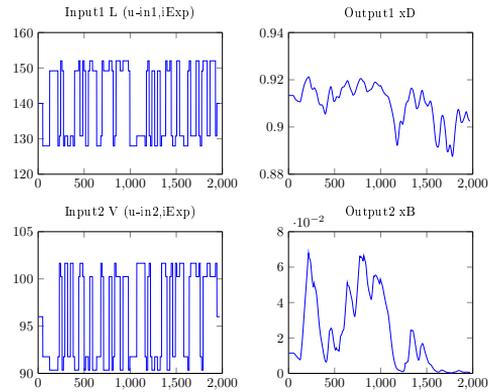
**Figure 6.4:** Simulator inputs and outputs (Exp. 1–6)
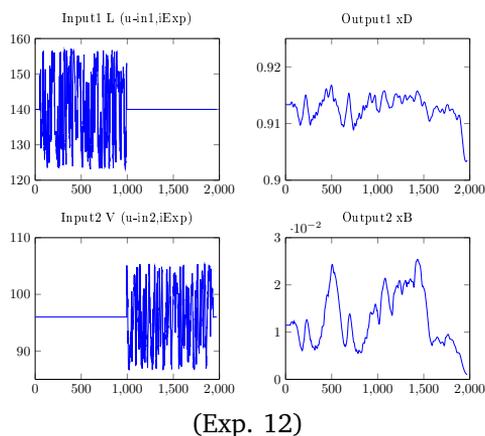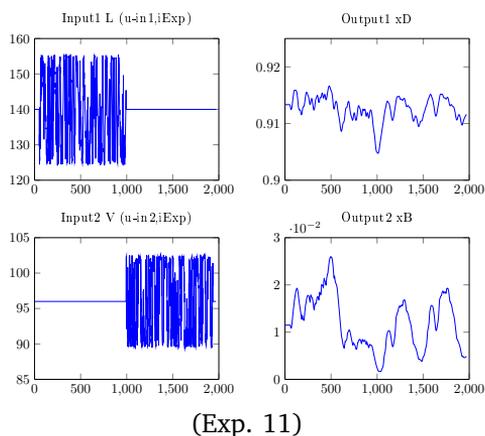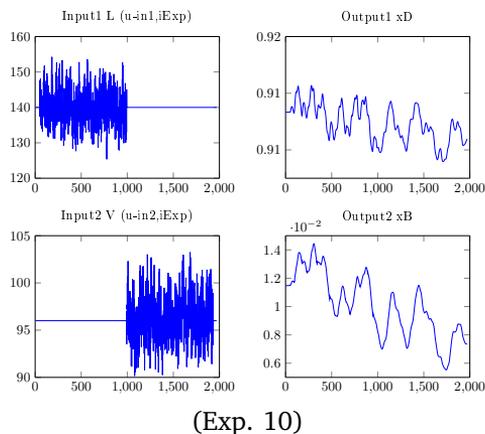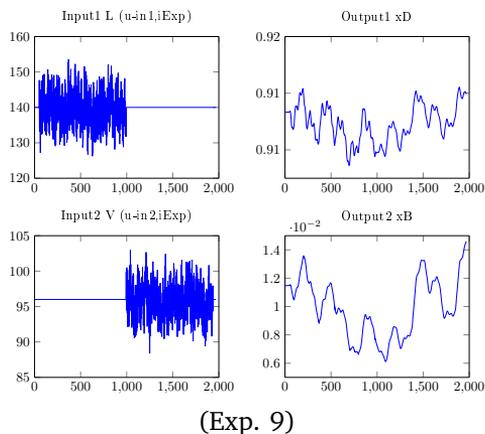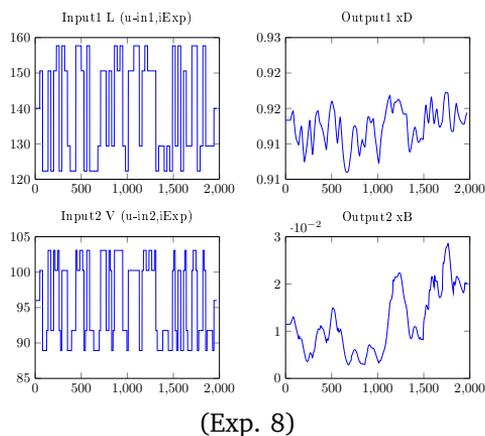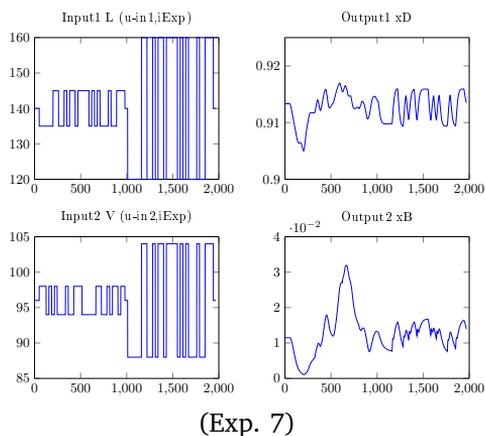
**Figure 6.5:** Simulator inputs and outputs (Exp. 7–12)
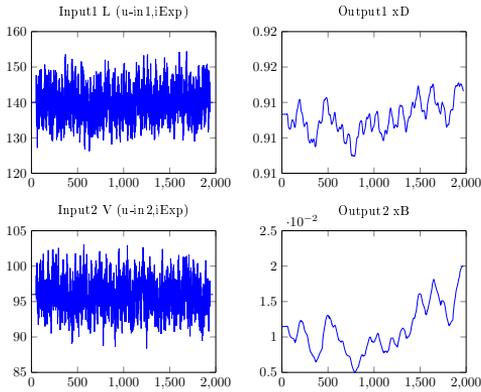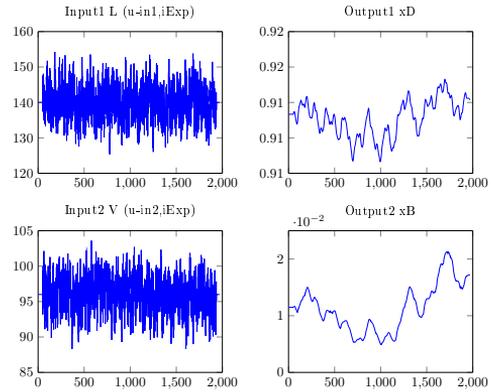
(Exp. 13)                                                       (Exp. 14)
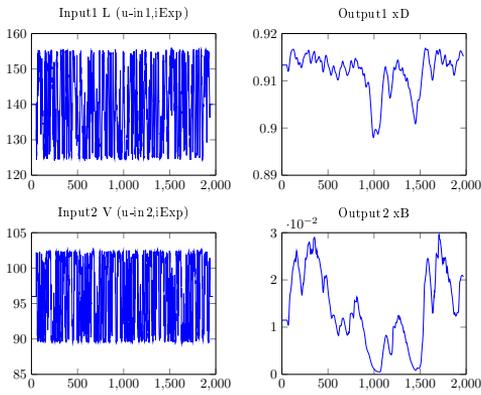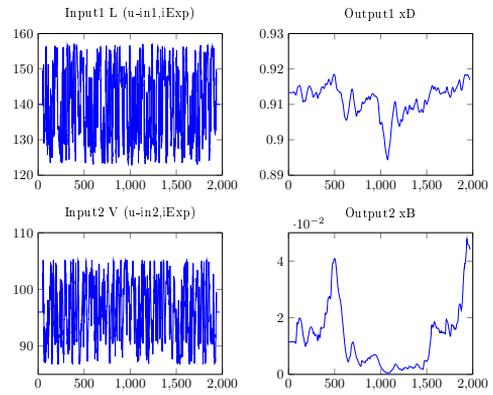
(Exp. 15)                                                       (Exp. 16)
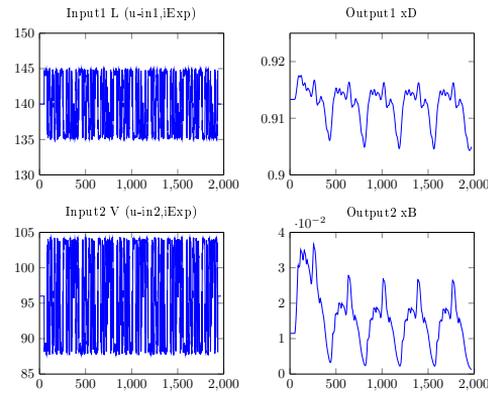
(Exp. 17)

**Figure 6.6:** Simulator inputs and outputs (Exp. 13–17)

## 6.3 Experiment design example - a $4 \times 4$ column stripper system

The performance of designed input excitations are compared on a column stripper system [Alatiqi and Luyben, 1986]. The system contains a broad range of time constants and time delays. It is a highly interactive and ill-conditioned (condition number > 100) system. The transfer function matrix of the system is

$$G(s) = \begin{bmatrix} \dfrac{4.09e^{-1.3s}}{(33s+1)(8.3s+1)} & \dfrac{6.36e^{-1.2s}}{(31.6s+1)(20s+1)} & \dfrac{-0.25e^{-1.4s}}{21s+1} & \dfrac{-0.49e^{-6s}}{22s+1} \\[2ex] \dfrac{-4.17e^{-5s}}{45s+1} & \dfrac{6.93e^{-1.02s}}{44.6s+1} & \dfrac{-0.05e^{-6s}}{(34.5s+1)^2} & \dfrac{1.53e^{-3.8s}}{48s+1} \\[2ex] \dfrac{1.73e^{-18s}}{(13s+1)^2} & \dfrac{5.11e^{-12s}}{(13.3s+1)^2} & \dfrac{4.61e^{-1.01s}}{18.5s+1} & \dfrac{-5.49e^{-1.5s}}{15s+1} \\[2ex] \dfrac{-11.2e^{-2.6s}}{(43s+1)(6.5s+1)} & \dfrac{14(10s+1)e^{-0.02s}}{(45s+1)(17.4s^2+3s+1)} & \dfrac{-0.1e^{-0.05s}}{(31.6s+1)(5s+1)} & \dfrac{4.49e^{-0.6s}}{(48s+1)(6.3s+1)} \end{bmatrix} \qquad (6.3.1)$$

A Gaussian distributed measurement noise $[s^2 = 0.03]$ was added to all outputs. A sampling time $T_s = 1$ min was used for all simulations. Various input design methods are used as summarized in Table 6.1. The step signals were designed with length of 330 min.

From the step changes, we have found that the system under study has dominating time constants $\tau_H = 50$ min and $\tau_L = 12$ min, whereas $\alpha$ and $\beta$ are chosen as 2 and 3, respectively. The parameter $\beta = 3$ signifies the 95% settling time of the system.

Now, with the help of 'cut and shifting technique' we can design least-correlated PRBS signal and apply simultaneously among all inputs with the help of following equations [Lee and Rivera, 2006]

$$T_{sw} \leq \frac{2.8 \ \tau_L}{\alpha} \qquad (6.3.2)$$

$$D \geq \frac{5 \ \tau_H}{T_{sw}} \qquad (6.3.3)$$

$$N_s \geq \max\left( \frac{2\pi \ \beta \ \tau_H}{T_{sw}}, \frac{5 \ p \ \tau_H}{T_{sw}} \right) \qquad (6.3.4)$$

$$N_s = 2^{n_r} - 1$$

For $n_r = 6$, we get $N_s = 2^6 - 1 = 63$ . Now, $T_{sw} \geq 15.87$ and $T_{sw} = 16$ minutes has chosen.

We found, the longest experiment would be the sequential PRBS inputs, where each four inputs are excited one after another. For this case, the experiment length ($T_L$) would

be 4032 minutes with $T_{sw} = 16$. However, for the experiments with simultaneous inputs (e.g. Exp. 4, prbsUnc), we can use longer PRBS sequences, for example, $n_r = 7$ or 8. Since $4032/16 = 252$ is closed to $N_s = 2^8 - 1 = 255$, we have chosen $n_r = 8$ for long PRBS sequences. All experiment signals were designed considering this (lengths) assumption. Figure 6.7 shows all inputs and outputs from one of the experiment is shown in. A switching time $T_{sw} = 16$ min, two PRBS length of 63 and 255 ($2^n - 1$; for $n$ number of bits) are considered for short (for Exp. 3, prbsSeq ) and long (for Exp. 4, prbsUnc ) PRBS, respectively. Total experiment length ( $T_L = 4080$ min) was the same for all cases.
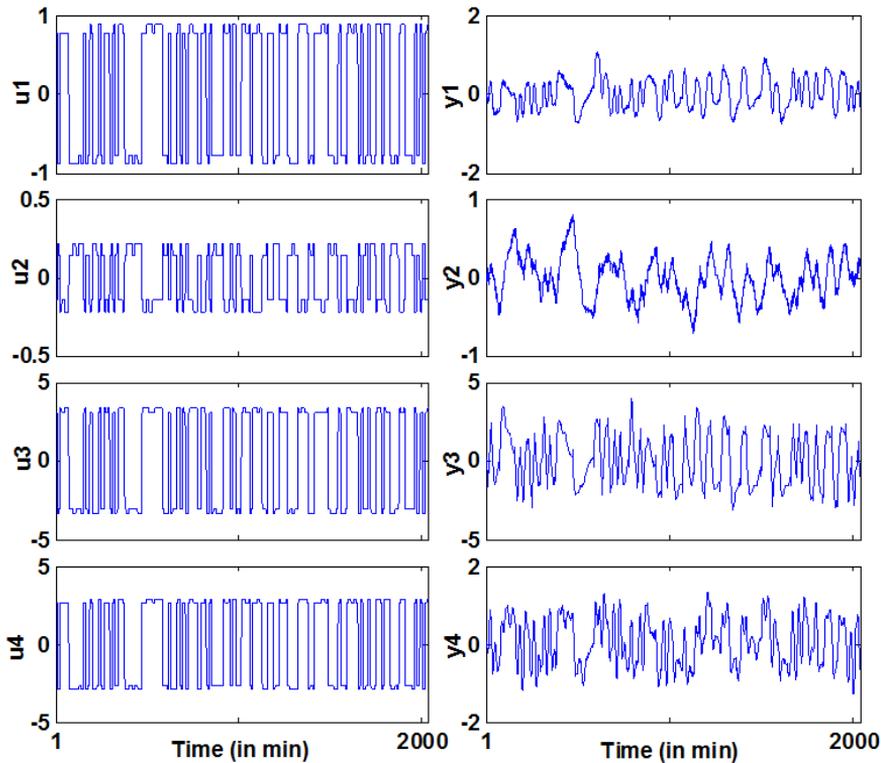


**Figure 6.7:** Input-output signals with simultaneous PRBS in all gain-directions (Exp. 8)

**Table 6.2:** Input signal characteristics

| Exp. | | Input CF($u_i$) | Avg. CF($u$) | PIPS($u_i$) % | Avg. PIPS ($u$) % | $min(u_i)$ | $max(u_i)$ | $max(|\Delta u_i|)$ |
|---|---|---|---|---|---|---|---|---|
| 1 stepSeq | $u_1$ | 2.487 | | 40.22 | | -0.094 | 0.094 | 0.188 |
| | $u_2$ | 2.487 | 2.487 | 40.22 | 40.22 | -0.068 | 0.068 | 0.136 |
| | $u_3$ | 2.487 | | 40.22 | | -0.26 | 0.260 | 0.520 |
| | $u_4$ | 2.487 | | 40.22 | | -0.165 | 0.165 | 0.330 |
| 2 stepDir | $u_1$ | 2.471 | | 40.47 | | -0.589 | 0.589 | 1.178 |
| | $u_2$ | 2.365 | 2.452 | 42.28 | 40.80 | -0.129 | 0.129 | 0.129 |
| | $u_3$ | 2.486 | | 40.23 | | -2.265 | 2.265 | 4.531 |
| | $u_4$ | 2.486 | | 40.23 | | -1.965 | 1.965 | 3.930 |
| 3 prbsSeq | $u_1$ | 2.011 | | 49.72 | | -0.141 | 0.141 | 0.282 |
| | $u_2$ | 2.012 | 2.012 | 49.70 | 49.70 | -0.102 | 0.102 | 0.204 |
| | $u_3$ | 2.012 | | 49.70 | | -0.39 | 0.390 | 0.780 |
| | $u_4$ | 2.012 | | 49.70 | | -0.248 | 0.248 | 0.495 |
| 4 prbsUnc | $u_1$ | 1.000 | | 100 | | -0.078 | 0.078 | 0.157 |
| | $u_2$ | 1.000 | 1.000 | 100 | 100.00 | -0.057 | 0.057 | 0.113 |
| | $u_3$ | 1.000 | | 100 | | -0.217 | 0.217 | 0.433 |
| | $u_4$ | 1.000 | | 100 | | -0.138 | 0.138 | 0.275 |
| 5 prbsSeqDir | $u_1$ | 2.000 | | 50.01 | | -0.589 | 0.589 | 1.178 |
| | $u_2$ | 1.914 | 1.984 | 52.26 | 50.42 | -0.129 | 0.129 | 0.258 |
| | $u_3$ | 2.011 | | 49.71 | | -2.265 | 2.265 | 4.531 |
| | $u_4$ | 2.011 | | 49.72 | | -1.965 | 1.965 | 3.930 |
| 6 prbsSimDir | $u_1$ | 1.162 | | 86.05 | | -0.551 | 0.551 | 0.943 |
| | $u_2$ | 1.481 | 1.179 | 67.53 | 86.62 | -0.161 | 0.161 | 0.206 |
| | $u_3$ | 1.032 | | 96.86 | | -1.872 | 1.872 | 3.625 |
| | $u_4$ | 1.041 | | 96.05 | | -1.637 | 1.637 | 3.144 |
| 7 prbsDirCorrSeq | $u_1$ | 1.402 | | 71.31 | | -0.589 | 0.589 | 1.178 |
| | $u_2$ | 1.294 | 1.379 | 77.18 | 72.58 | -0.129 | 0.129 | 0.258 |
| | $u_3$ | 1.408 | | 71.00 | | -2.265 | 2.265 | 4.531 |
| | $u_4$ | 1.411 | | 70.85 | | -1.965 | 1.965 | 3.930 |
| 8 prbsDirCorrSim | $u_1$ | 1.124 | | 88.98 | | -0.472 | 0.472 | 0.833 |
| | $u_2$ | 1.316 | 1.127 | 75.97 | 89.65 | -0.131 | 0.131 | 0.182 |
| | $u_3$ | 1.000 | | 100.00 | | -1.755 | 1.755 | 3.510 |
| | $u_4$ | 1.068 | | 93.66 | | -1.487 | 1.487 | 2.779 |
| 9 msSrdSeq | $u_1$ | 5.295 | | 18.91 | | -0.531 | 0.532 | 0.559 |
| | $u_2$ | 6.202 | 5.685 | 17.32 | 18.34 | -0.39 | 0.454 | 0.417 |
| | $u_3$ | 5.549 | | 19.23 | | -1.536 | 1.341 | 1.508 |
| | $u_4$ | 5.694 | | 17.89 | | -1.01 | 0.973 | 0.993 |
| 10 msSrdSeqDir | $u_1$ | 5.659 | | 18.00 | | -2.164 | 2.085 | 2.126 |
| | $u_2$ | 5.419 | 5.616 | 18.80 | 18.15 | -0.456 | 0.474 | 0.441 |
| | $u_3$ | 5.692 | | 17.89 | | -8.323 | 8.019 | 8.177 |
| | $u_4$ | 5.692 | | 17.90 | | -7.22 | 6.955 | 7.093 |
| 11 msGlmSeq | $u_1$ | 3.073 | | 32.70 | | -0.21 | 0.207 | 0.400 |
| | $u_2$ | 2.497 | 2.784 | 40.16 | 36.42 | -0.124 | 0.123 | 0.241 |
| | $u_3$ | 3.042 | | 33.13 | | -0.577 | 0.568 | 1.092 |
| | $u_4$ | 2.524 | | 39.68 | | -0.305 | 0.304 | 0.563 |
| 12 msGlmSeqDir | $u_1$ | 2.509 | | 39.92 | | -0.677 | 0.675 | 1.251 |
| | $u_2$ | 2.401 | 2.489 | 41.70 | 40.25 | -0.148 | 0.148 | 0.263 |
| | $u_3$ | 2.524 | | 39.69 | | -2.603 | 2.594 | 4.812 |
| | $u_4$ | 2.523 | | 39.69 | | -2.258 | 2.250 | 4.174 |
| 13 msSrdSim | $u_1$ | 2.794 | | 35.79 | | -0.311 | 0.311 | 0.312 |
| | $u_2$ | 3.102 | 2.901 | 34.68 | 35.47 | -0.215 | 0.250 | 0.229 |
| | $u_3$ | 2.750 | | 36.36 | | -0.845 | 0.845 | 0.876 |
| | $u_4$ | 2.958 | | 35.07 | | -0.577 | 0.535 | 0.546 |
| 14 msSrdSimDir | $u_1$ | 3.006 | | 33.63 | | -1.379 | 1.349 | 1.363 |
| | $u_2$ | 3.171 | 3.027 | 32.93 | 34.08 | -0.305 | 0.333 | 0.310 |
| | $u_3$ | 2.961 | | 34.79 | | -5.193 | 4.889 | 4.964 |
| | $u_4$ | 2.969 | | 34.97 | | -4.517 | 4.185 | 4.259 |
| 15 msGlmSim | $u_1$ | 1.537 | | 65.07 | | -0.11 | 0.110 | 0.207 |
| | $u_2$ | 1.271 | 1.397 | 79.69 | 72.51 | -0.066 | 0.064 | 0.125 |
| | $u_3$ | 1.518 | | 65.89 | | -0.299 | 0.299 | 0.566 |
| | $u_4$ | 1.261 | | 79.38 | | -0.158 | 0.158 | 0.292 |
| 16 msGlmSimDir | $u_1$ | 1.442 | | 69.48 | | -0.445 | 0.443 | 0.720 |
| | $u_2$ | 1.879 | 1.480 | 53.96 | 69.47 | -0.133 | 0.129 | 0.160 |
| | $u_3$ | 1.289 | | 77.74 | | -1.52 | 1.515 | 2.754 |
| | $u_4$ | 1.309 | | 76.68 | | -1.329 | 1.340 | 2.387 |
| 17 msGlmZipm | $u_1$ | 1.279 | | 78.63 | | -0.718 | 0.710 | 1.267 |
| | $u_2$ | 1.465 | 1.318 | 70.14 | 76.96 | -0.171 | 0.181 | 0.265 |
| | $u_3$ | 1.264 | | 79.51 | | -2.726 | 2.700 | 4.866 |
| | $u_4$ | 1.262 | | 79.58 | | -2.362 | 2.341 | 4.222 |

### 6.3.1   Plant friendly parameters

The plant friendly parameters of the designed input signals of all the experiments are calculated and compared. The crest factor (CF), PIPS and their average values over all four inputs are summarized in Table 6.2. For comparison of various inputs we also have considered the maximum ($\max(u_i)$) and minimum ($\min(u_i)$) value of the input as well as the difference (or change) between two consecutive input samples ($|\Delta u_i|$). It was found that the PRBS signal in Exp. 4, (prbsUnc) has the smallest average CF (= 1), while Exp. 9 (msSrdSeq), has the highest CF (=5.685).
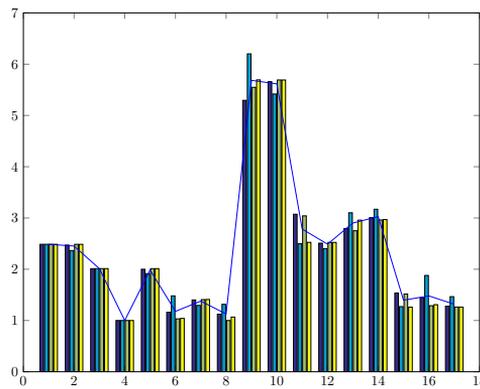


**Figure 6.8:** Crest factor of the design inputs and their average values (—)
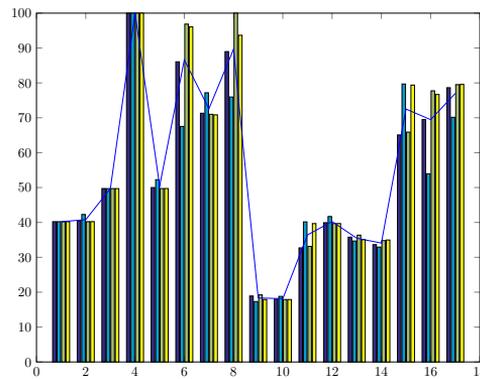


**Figure 6.9:** PIPS of the design inputs and their average values (—)

The determinant of corr($Y$) ) value is calculated based on the output data for each experiment. This scalar value can provide an indication if the designed inputs are able to extract various gain direction informations equally well or not. Figure 6.10 shows $\det(corr(Y))$ values of each designed experiments in $4 \times 4$ system. A few design methods
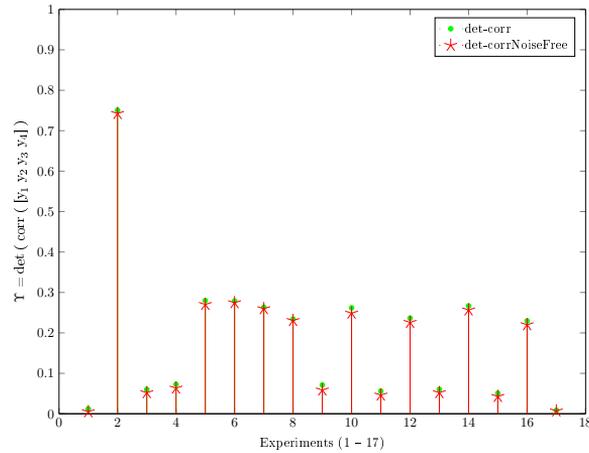
**Figure 6.10:** Determinant of output-correlation ($\Upsilon$) of designed experiments in $4 \times 4$ system

are able to extract all gain directions (higher values), whereas a few are not able to extract so well (smaller values).

### 6.3.2   Model fitting and cross-validation

First, a model is determined from the experimental data. The average fit percentage value of all four outputs of the system is given in Table 6.3. Moreover, the performance of the estimated model has been validated through the other experiments. From Figures 6.11 and 6.12, we can clearly see some differences among the models generated from the data. We found that models from the experiments that considered the gain directions during the input signal design can perform better in cross validation in all cases (best and worst cases).

Experiment 14 (msSrdSimDir) gave best model based on cross validation [best fit: 91.83% and worst fit: 83.14%]. Experiment 10 (msSrdSeqDir) [best 91.67% and worst 84.55%], Exp. 12 (msGlmSeqDir) [best 91.81% and worst 82.63%], Exp. 16 (msGlm-SimDir) [best 91.78% and worst 82.27%] and Exp. 17 (msGlmZipm) [best 91.53% and worst 82.87%] also gave near good models based on cross validation. In all these experiments multisine excitation was used and directional information was considered in the design process. Hence, it is important to consider directional information in the input design. On the other hand, experiment 2 (stepDir), gave a good model with the best fir to identification data [best 92.63%] but the model did not perform well in cross validation [worst 82.63%]. However, as explained earlier, steps are very useful for estimation of basic system properties. PRBS excitations with gain directions, Exp. 5 (prbsSeqDir)

**Table 6.3:** Overall performance of the models

| Model | Input excitation | Average model-fit (%) | Average cross-fit (%) | Best cross-fit (%) | Worst cross-fit (%) |
|-------|------------------|------------------------|------------------------|--------------------|---------------------|
| 1 | stepSeq | 92.72 | 79.24 | 89.55 | 64.51 |
| 2 | stepDir | 88.02 | 84.94 | 92.63 | 73.56 |
| 3 | prbsSeq | 89.36 | 81.94 | 91.54 | 72.49 |
| 4 | prbsUnc | 89.26 | 83.95 | 91.35 | 75.52 |
| 5 | prbsSeqDir | 85.21 | 87.28 | 91.08 | 82.08 |
| 6 | prbsSimDir | 88.78 | 87.10 | 91.11 | 82.35 |
| 7 | prbsDirCorrSeq | 88.79 | 87.13 | 91.32 | 81.66 |
| 8 | prbsDirCorrSim | 89.32 | 87.24 | 91.37 | 81.31 |
| 9 | msSrdSeq | 88.12 | 83.30 | 91.44 | 75.65 |
| 10 | msSrdSeqDir | 83.38 | 87.78 | 91.67 | 84.55 |
| 11 | msGlmSeq | 89.99 | 83.31 | 91.64 | 74.72 |
| 12 | msGlmSeqDir | 85.85 | 87.69 | 91.81 | 82.63 |
| 13 | msSrdSim | 89.34 | 83.61 | 91.54 | 75.99 |
| 14 | msSrdSimDir | 85.21 | 87.81 | 91.83 | 83.14 |
| 15 | msGlmSim | 90.41 | 82.76 | 91.71 | 73.75 |
| 16 | msGlmSimDir | 87.21 | 87.50 | 91.78 | 82.27 |
| 17 | msGlmZipm | 92.15 | 87.44 | 91.53 | 82.87 |

and Exp. 6 (prbsSimDir), perform significantly better in cross validation [worst $\approx 82\%$] compared to their counterparts with no directional information, Exp. 3 (prbsSeq) and Exp. 4 (prbsUnc) [worst $\approx 72 - 75\%$], as shown in Table 6.3. The design methods with combination of highly correlated and uncorrelated signals, Exp. 7 (prbsDirCorrSeq) and Exp. 8 (prbsDirCorrSim), did not perform [worst $\approx 81\%$] as well as PRBS with gain directions (Exp. 5 & 6). However, they perform better than standard PRBS without gain direction information (Exp. 3 & Exp. 4), as anticipated.

### 6.3.3   Singular values of estimated models

The singular values ($\sigma_i$) of the steady-state gain matrix of the generated models are shown in Table 6.4. Many models are able to produce condition numbers near to the true ones, except Exp. 4 (prbsUnc), Exp. 9 (msSrdSeq), Exp. 11 (msGlmSeq), Exp. 13 (msSrdSim) and Exp. 15 (msGlmSim), where estimated values are at-least 16% lower. Whereas in case of $\sigma_i$ values, it is quite clear that the only some models are able to capture the singular values near to its true values. One reason behind this could be not enough excitation in weaker directions of the system.

**Table 6.4:** Singular Values ($\sigma_i$) and their ratios ($\xi_i$) of the estimated models

| Model from | | Singular values | | | | Ratios | | |
| | | $\sigma_1$ | $\sigma_2$ | $\sigma_3$ | $\sigma_4$ | $\xi_1 = \frac{\sigma_1}{\sigma_2}$ | $\xi_2 = \frac{\sigma_1}{\sigma_3}$ | $\xi_3{}^* = \frac{\sigma_1}{\sigma_4}$ |
|---|---|---|---|---|---|---|---|---|
| Exp. 1 | stepSeq | 20.30 | 10.19 | 5.05 | 0.17 | 2.0 | 4.0 | 121.07 |
| Exp. 2 | stepDir | 20.31 | 10.14 | 5.05 | 0.16 | 2.0 | 4.0 | 125.03 |
| Exp. 3 | prbsSeq | 19.34 | 10.32 | 5.08 | 0.14 | 1.9 | 3.8 | 138.03 |
| Exp. 4 | prbsUnc | 18.96 | 10.35 | 5.09 | 0.18 | 1.8 | 3.7 | 103.31 |
| Exp. 5 | prbsSeqDir | 18.65 | 10.20 | 4.98 | 0.15 | 1.8 | 3.7 | 122.53 |
| Exp. 6 | prbsSimDir | 17.90 | 10.16 | 5.03 | 0.15 | 1.8 | 3.6 | 119.79 |
| Exp. 7 | prbsDirCorrSeq | 17.38 | 10.49 | 5.10 | 0.16 | 1.7 | 3.4 | 110.20 |
| Exp. 8 | prbsDirCorrSim | 18.28 | 10.38 | 5.10 | 0.16 | 1.8 | 3.6 | 115.19 |
| Exp. 9 | msSrdSeq | 18.89 | 10.35 | 5.08 | 0.20 | 1.8 | 3.7 | 96.25 |
| Exp.10 | msSrdSeqDir | 18.51 | 10.22 | 4.99 | 0.15 | 1.8 | 3.7 | 122.34 |
| Exp.11 | msGlmSeq | 19.07 | 10.34 | 5.07 | 0.19 | 1.8 | 3.8 | 102.35 |
| Exp.12 | msGlmSeqDir | 18.52 | 10.08 | 5.11 | 0.16 | 1.8 | 3.6 | 118.85 |
| Exp.13 | msSrdSim | 19.06 | 10.37 | 5.09 | 0.18 | 1.8 | 3.7 | 105.08 |
| Exp.14 | msSrdSimDir | 18.90 | 10.31 | 5.11 | 0.15 | 1.8 | 3.7 | 122.64 |
| Exp.15 | msGlmSim | 19.14 | 10.35 | 5.06 | 0.20 | 1.8 | 3.8 | 96.22 |
| Exp.16 | msGlmSimDir | 18.85 | 10.10 | 5.05 | 0.16 | 1.9 | 3.7 | 120.20 |
| Exp.17 | msGlmZipm | 19.30 | 10.37 | 5.17 | 0.16 | 1.9 | 3.7 | 121.68 |
| TRUE system | – | 20.23 | 10.11 | 5.05 | 0.16 | 2.0 | 4.0 | 125.2 |

$\xi_3{}^*$ also known as the condition number (CN) of the system.

The singular values of the identified models and the true system in frequency domain are shown in Figure 6.11. Most models actually fit of the singular values with the true system quite well for lower frequencies, but less so for higher frequencies. This is expected as the input signals were designed primarily for the lower range of the frequencies, (for PRBS $\omega_L = 0.01$ , $\omega_H = 0.2$), and we fit a lower order model [Eq. 6.1.1] of the true system.

In case of the fourth singular value ($\sigma_4$), the mismatch between estimated model and true system is substantial. One reason of this mismatch could be system dynamics mismatch in true system compare to the estimated lower order model. Also, note that the amplitude of $\sigma_4$ in all frequencies is quite low (below -15dB to as low as -60dB) [Figure 6.11].
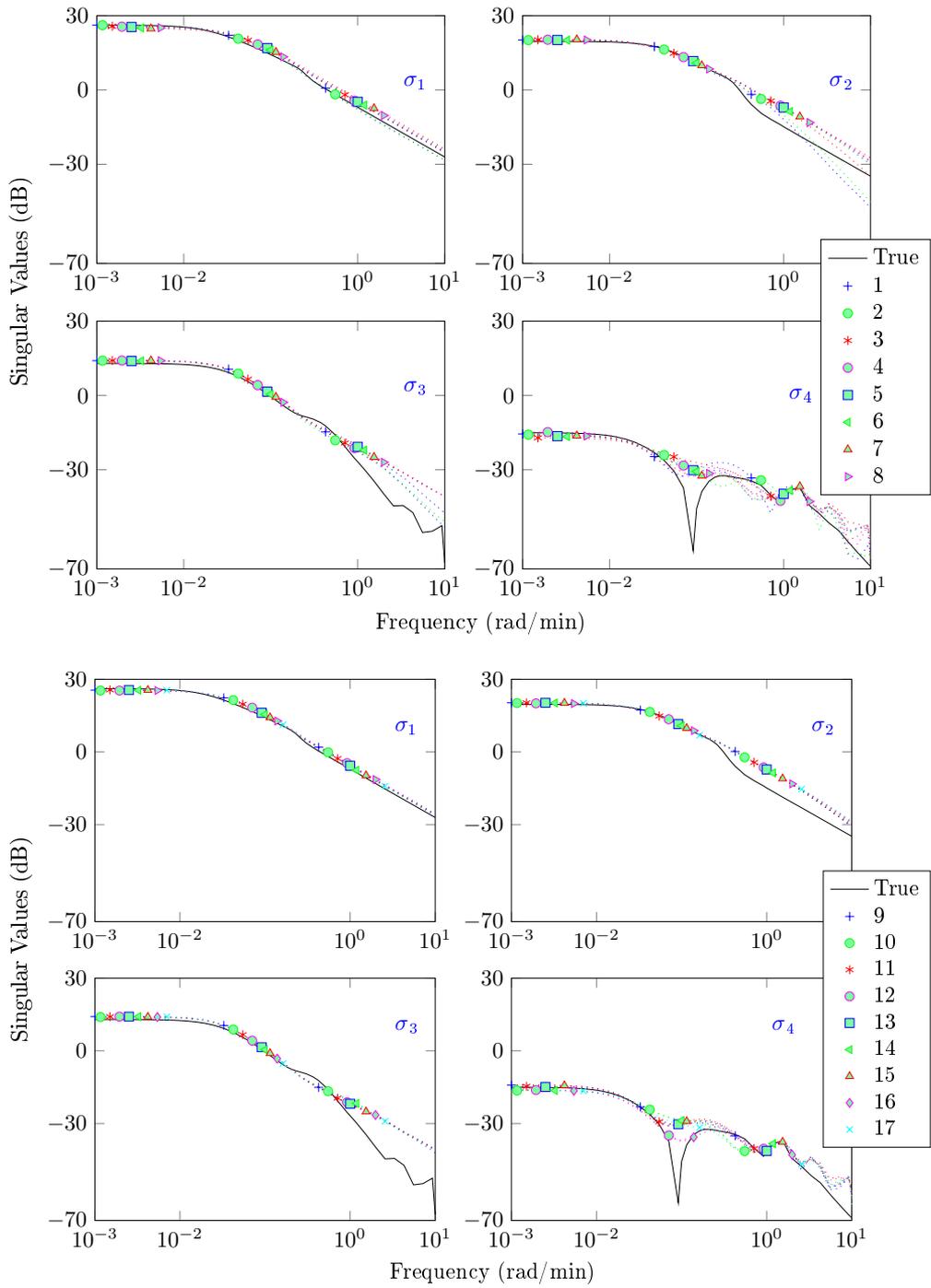
**Figure 6.11:** Singular values in frequency domain (a comparison with true system and estimated models)

### 6.3.4  Projections of output data on gain directions

The projections on gain directions of the outputs are obtained using the $W$ matrix from the SVD (Eq. 5.2.1). This projection of all experiments onto the low-gain and high-gain spaces are shown in Figure 6.12. Each sub-figure illustrates data-set from different design methods and how well they are balanced or imbalanced in strongest ($\sigma_1$) to weakest ($\sigma_4$) direction.

Using Eq. 5.5.2, the projection of output along the high-gain direction is given by $y_1^\times(k) = w_1^T y(k)$ and that in low-gain direction is $y_4^\times(k) = w_4^T y(k)$. In Table 6.5, the standard deviations of projections onto all four spaces along the gain-directions are presented.

**Table 6.5:** Standard deviation of projection along gain directions

| Exp. | Input excitation | High-gain | Middle-gain 1 | Middle-gain 2 | Low-gain | *min/max* Ratio |
|------|------------------|-----------|---------------|---------------|----------|-----------------|
| 1  | stepSeq         | 0.603 | 0.555 | 0.349 | 0.018 | 0.030 |
| 2  | stepDir         | 0.252 | 0.268 | 0.204 | 0.179 | 0.668 |
| 3  | prbsSeq         | 0.568 | 0.605 | 0.409 | 0.052 | 0.086 |
| 4  | prbsUnc         | 0.573 | 0.672 | 0.443 | 0.058 | 0.086 |
| 5  | prbsSeqDir      | 0.275 | 0.457 | 0.236 | 0.123 | 0.269 |
| 6  | prbsSimDir      | 0.453 | 0.743 | 0.383 | 0.193 | 0.260 |
| 7  | prbsDirCorrSeq* | 0.483 | 0.733 | 0.411 | 0.184 | 0.251 |
| 8  | prbsDirCorrSim* | 0.520 | 0.603 | 0.302 | 0.174 | 0.289 |
| 9  | msSrdSeq        | 0.470 | 0.562 | 0.368 | 0.045 | 0.080 |
| 10 | msSrdSeqDir     | 0.231 | 0.421 | 0.222 | 0.100 | 0.238 |
| 11 | msGlmSeq        | 0.511 | 0.618 | 0.413 | 0.044 | 0.071 |
| 12 | msGlmSeqDir     | 0.254 | 0.451 | 0.232 | 0.107 | 0.237 |
| 13 | msSrdSim        | 0.547 | 0.635 | 0.415 | 0.050 | 0.079 |
| 14 | msSrdSimDir     | 0.285 | 0.510 | 0.273 | 0.122 | 0.239 |
| 15 | msGlmSim        | 0.550 | 0.650 | 0.428 | 0.046 | 0.071 |
| 16 | msGlmSimDir     | 0.295 | 0.525 | 0.268 | 0.123 | 0.234 |
| 17 | msGlmZipm*      | 0.487 | 0.905 | 0.419 | 0.221 | 0.244 |

* only low-gain direction information are used in input design.

Based on the *min/max* ratio of the standard deviation of projections, we can categorize the experiments in Table 6.5 into three major groups: (a) $> 0.5$, (b) in the range of $0.2$ to $0.3$, and (c) $< 0.1$ . Experiment 2 (stepDir, $0.668$) is the only experiment belongs to group (a). Group (b) contains nine experiments (Exp. 5, 6, 7, 8, 10, 12, 14, 16 and 17) and the rest (Exp. 1, 3, 4, 9, 11, 13, 15) belong to group (c), which fail to generate well-balanced data in all gain-directions. This ratio is one of the indication, how well the

system is excited in various gain directions i.e. in terms of gain directions collected data
are balanced or not.

The data points are scattered different ways in low- and high-gain direction in Figure 6.12. The reason why the standard deviations are not equal in all directions in the
cases that is designed to be in balance, is that the projection data points are calculated
based on dynamic output data, though the design methods concerned steady-state only.
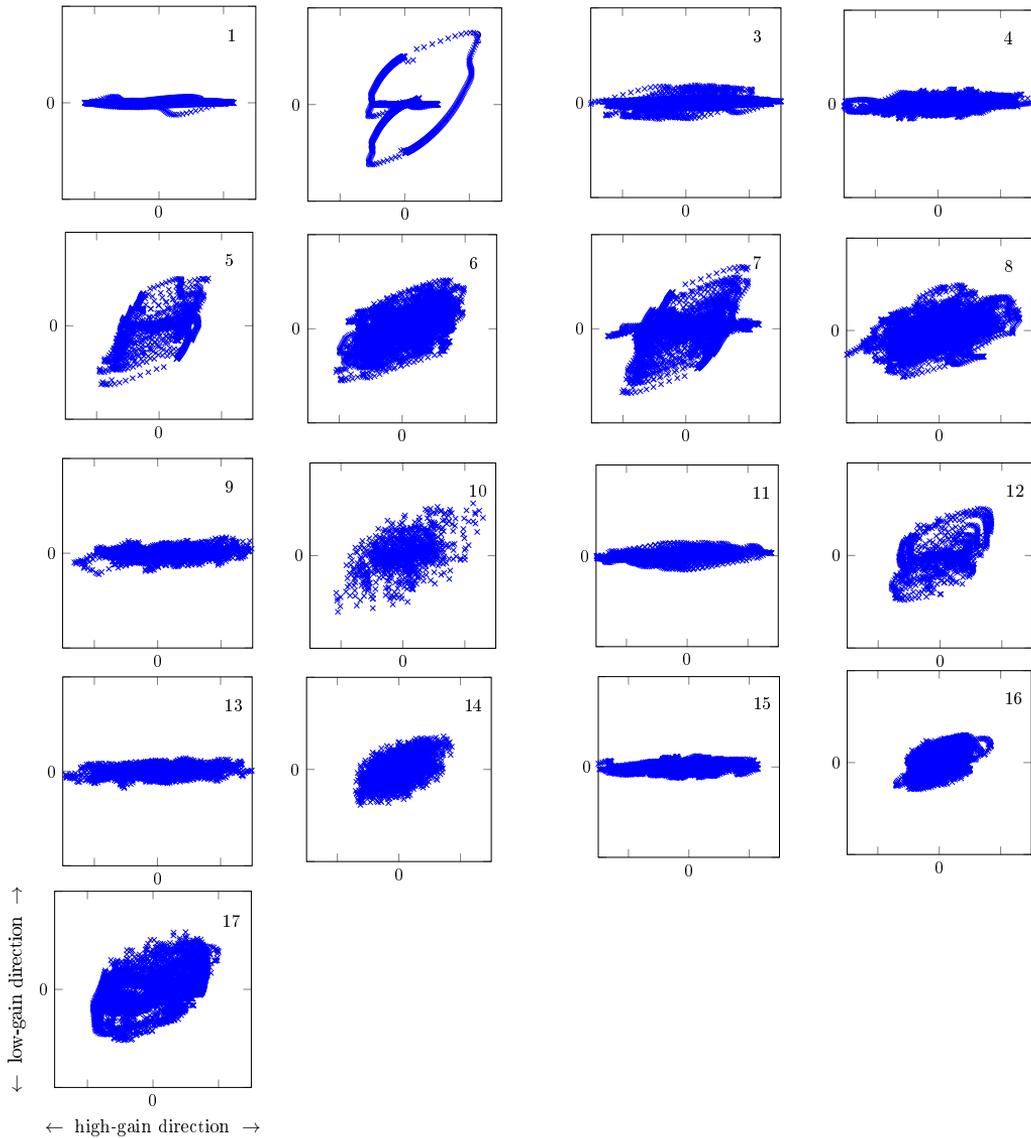


**Figure 6.12:** Output data projections for all experiments (low-gain vs. high-gain direction)

## 6.4 Ill conditioned system and its $\kappa, \sigma$ values

The condition number ($\kappa$), from which the degree of ill-conditioning is inferred, is the ratio between the highest singular value ($\sigma_1$) and the smallest singular value ($\sigma_n$). However if $n > 2$ the system might also be ill-conditioned due to other small singular values. Consider the transfer function matrix

$$G(s) = \begin{bmatrix} \frac{0.173e^{-3.79s}}{(21.74s+1)^2} & \frac{-3.4e^{-7.75s}}{(22.22s+1)^2} & \frac{-8.66e^{-1.59s}}{11.36s+1} & \frac{11.94e^{-27.33s}}{33.3s+1} \\ \frac{-0.268e^{-60s}}{400s+1} & \frac{2.79e^{-0.71s}}{(66.67s+1)^2} & \frac{5.05e^{-2.24s}}{14.29s+1} & \frac{-5.65e^{-8.72s}}{(250s+1)^2} \\ \frac{-0.084e^{-0.68s}}{(2.38s+1)^2} & \frac{0.94e^{-0.59s}}{(7.14s+1)^2} & \frac{1.31e^{-0.42s}}{(1.43s+1)^2} & -0.56e^{-s} \\ \frac{-0.047e^{-0.52s}}{11.11s+1} & \frac{-15.5e^{-0.48s}}{(6.90s+1)^2} & \frac{-9.03e^{-1.91s}}{12.19s+1} & -8.71e^{-s} \end{bmatrix} \tag{6.4.1}$$

with singular values, $diag(\Sigma) = [20.39, 16.71, 0.185, 0.156]$. In this example, even after compensating lowest gain-direction, system will remain ill-conditioned due to the second smallest singular value. Therefore, exciting only the weakest gain direction (which is popular in dealt with $2 \times 2$ systems) is not enough here. One need to consider all gain directions, or at-least the potential ones.

Higher gain direction dominates in normal excitation. Hence, one has to excite all those directions with $\xi_i \gg 1$, where $\xi_i = \sigma_1/\sigma_i$ .

# Conclusion

> *"Now this is not the end. It is not even the beginning of the end. But it is, perhaps, the end of the beginning."*
>
> *- W. Churchill*

In this chapter, we present a brief summary of the results. We also mention some of the important problems that will be addressed in the near future.

## 7.1   Summary of results in the thesis

We have discussed and compared various input design methods (with general signals like step, PRBS, multisine ) in order to find the best method to tackle the underlying challenge in ill-conditioned system identification. The plant friendly input aspect is also discussed and considered in our design. We have used a few methods (such as projection of output data along the gain directions, determinant of correlation matrix of outputs etc.) to evaluate various methods in order to design balanced outputs in directional systems.

Prior to the input design, in addition to other practical checks, one should check if the underlying system is *ill-conditioned or not?* This can be done using step tests. Using the ratios of singular values of a system one can identify all weak gain directions, which need special attention in order to excite the system in balance. The results indicate that in order

to obtain a model suitable for control design, it is advisable not only to excite the low-gain direction, but also excite inputs according to all weak ones if not all gain directions. Since an accurate steady-state system model is unavailable in most situations, one can use an approximate model, which may be acquired from step test to obtain estimates of $V_i$ and $\sigma_i$.

Two parameters have been defined to check whether the collected data are balanced or not. These parameters are *the scattering of projections along the gain directions $y^*$* and determinant of output-correlation ($\Upsilon$). Both of these parameters only need the output data of the system. Hence, one can use for any multivariable system ($n \times n, n > 2$) and easily determine whether the data are in good balance or not, before the modeling process. This is useful in ill-conditioned system modeling.

## 7.2   Future work

Following areas might be interesting to consider for continue research in near future.

1. **Input design for closed-loop MIMO identification**

   We have focused our work on open loop identification. However, it would be natural to extend the methods to closed loop identification.

2. **Practical implementation of input design**

   Collection of normal process data is quite common in industrial application. Much of them can not be used for system identification, however, using proper filtering algorithms some useful information of the system might be extracted. It would be interesting to study for suitable algorithms (may be some adaptiveness i the algorithm) to detect if the system is prone to ill-conditioning or not, using available process data. Accordingly, we could implement the necessary framework in system identification process.

3. **Input design for nonlinear systems**

   More study is required to consider nonlinear systems and the design of plant-friendly input signals. Multisine signals (e.g., odd and even) in frequency domain identification can be use for this purpose.

4. **Integration of model based control and input design**

Model based control (e.g. MPC, economic MPC etc.) is quite popular in industry. Generating a model based on process data is also a popular. Hence, integration of model based control with input design would be an interesting area to look upon. A number of model quality constraints can be formulated as linear functions of the covariance matrix in parameter estimation.

5. **Integration of optimal input design and robust control**

Input design can be formulated as an optimization problem, it would also be interesting to study optimal input design and robust control together.

# Bibliography

I. Alatiqi and W. Luyben. Control of a complex sidestream column/stripper distillation configuration. *Industrial & Engineering Chemistry, Process Design and Development*, 25(3):762–767, 1986. [55, 90]

A. C. Antoulas and B. D. O. Anderson. On the choice of inputs in identification for robust control. *Automatica*, 35(6):1009–1031, 1999. [22]

H. Barker and K. Godfrey. System identification with multi-level periodic perturbation signals. *Control Engineering Practice*, 7(6):717–726, 1999. [22]

I. Cameron and K. Hangos. *Process Modelling and Model Analysis*, volume 4. Academic Press, 2001. [5]

J. Conner and D. Seborg. An evaluation of mimo input designs for process identification. *Industrial & Engineering Chemistry Research*, 43(14):3847–3854, 2004. [39, 40, 41]

S. Gaikwad and D. Rivera. Control-relevant input signal design for multivariable system identification: Application to high-purity distillation. In *Proceedings of the 13th IFAC World Congress*, volume 1000, pages 349–354. Citeseer, 1996. [46]

M. Gevers. Identification for control: From the early achievements to the revival of experiment design. *European Journal of Control*, 11(4):335–352, 2005. [14]

R. Ghosh. Construction of plant friendly input excitations for ill-conditioned multivariable process identification. In *6th International Symposium on Communications, Control and Signal Processing (ISCCSP)*, pages 368–371, Athens, Greece, 2014. IEEE. [31, 90]

R. Ghosh, K. Häggblom, and J. Böling. Control-relevant input excitation for system identification of ill-conditioned $n \times n$ systems with $n > 2$. In *In Proc. 19th IFAC World Congress*, volume 19, pages 9382–9387, Cape Town, South Africa, 2014. [90]

R. Ghosh, K. Häggblom, and J. Böling. Input design for open-loop identification of ill-conditioned $n \times n$ systems with $n > 2$. *Submitted*, 2016. [32, 47]

K. Godfrey. Comparison of perturbation signals for linear system identification in the frequency domain. *IEE Proc. Control Theory Appl.*, 146(6):535–548, 1999. [22]

K. Godfrey, A. Tan, H. Barker, and B. Chong. A survey of readily accessible perturbation signals for system identification in the frequency domain. *Control Engineering Practice*, 13(11):1391–1402, 2005. [22, 23]

P. Guillaume, J. Schoukens, and R. Pintelon. Crest-factor minimization using nonlinear chebyshev approximation methods. *IEEE Transactions on Instrumentation and Measurement*, 40:982–989, 1991. [16, 20, 22, 23, 25, 27, 33, 42, 45, 48]

P. Guillaume, J. Schoukens, R. Pintelon, I. Kollar, H. Barker, A. Tan, and K. Godfrey. Wiener models of direction-dependent dynamic systems. *Automatica*, 39(1):127–133, 2003. [22, 42]

K. Häggblom. Modeling of flow dynamics for control of distillation columns. In *American Control Conference*, pages 785–790, Boston, USA, 1991. [80]

K. Häggblom and J. Böling. Multimodel identification for control of an ill-conditioned distillation column. *Journal of Process Control*, (427):2–5, 1998. [36, 37, 40, 85]

K. Häggblom and J. Böling. Experimental evaluation of input designs for multiple-input multiple-output open-loop system identification. In *Proc. IASTED Int. Conf. Control and Applications*, pages 157–163, Honolulu, USA, 2013. [37]

K. Häggblom and R. Ghosh. Tuning a distillation column simulator. In E.N. Pistikopoulos, M.C. Georgiadis, and A.C. Kokossis, editors, *21st European Symposium on Computer Aided Process Engineering*, volume 29 of *Computer Aided Chemical Engineering*, pages 823–827. Elsevier, 2011. [9]

K. Häggblom and R. Ghosh. Experiment designs for control-oriented mimo identification. In *Finnish Society of Automation Publication Series no. 44*, 44, Helsinki, Finland, 2015. Suomen Automaatioseuran. [32, 90]

R.G. Hakvoort. *System identification for rubust process control. Nominal models and error bounds*. PhD thesis, 1994. [35]

B. Huang, S. Ding, and Qin S. Closed-loop subspace identification: an orthogonal projection approach. *Journal of Process Control*, 15:53–66, 2005. [39]

A. J. Hugo. Process controller performance monitoring and assessment. *Control Arts Inc.*, 2001. [21]

E. Jacobsen and S. Skogestad. Inconsistencies in dynamic models for ill-conditioned plants: application to low-order models of distillation columns. *Industrial & Engineering Chemistry Research*, 33(3):631–640, 1994. [38, 41, 89]

T. Katayama, T. McKelvey, A. Sano, C. Cassandras, and M. Campi. Trends in systems and signals: Status report prepared by the ifac coordinating committee on systems and signals. *Annual Reviews in Control*, 30(1):5–17, 2006. [14]

C. Koung and J. MacGregor. Design of identification experiments for robust control. a geometric approach for bivariate processes. *Industrial & Engineering Chemistry Research*, 32(8):1658–1666, 1993. [16, 39, 40, 41]

C. Koung and J. MacGregor. Identification for robust multivariable control: the design of experiments. *Automatica*, 30:1541–1554, 1994. [38, 90]

H. Lee. *A Plant-Friendly Multivariable System Identification Framework Based on Identification Test Monitoring*. PhD thesis, Arizona State University, 2006. [15, 17, 25, 29, 30, 31, 33, 40, 41, 42, 45, 47, 48]

H. Lee and D. Rivera. Cr-ident: A matlab toolbox for multivariable control-relevant system identification. In *Proc. 14th IFAC Symposium on System Identification (SYSID)*, pages 708–713, Newcastle, Australia, 2006. [15, 31, 40, 41, 55]

W. Li and J. Lee. Control relevant identification of ill-conditioned systems: Estimation of gain directionality. *Computers & chemical engineering*, 20(8):1023–1042, 1996. [38, 90]

L. Ljung. *System Identification: Theory for the User*. Prentice-Hall, Inc., 1999. [5, 6, 11, 17, 18, 19, 49]

L. Ljung. *System Identification Toolbox Users Guide 2012b*, September 2012. [49]

Q. Meng, F. Fang, and J. Liu. Minimum-information-entropy-based control performance assessment. *Entropy*, 15(3):943–959, 2013. [21]

A. Micchi and G. Pannocchia. Comparison of input signals in subspace identification of multivariable ill-conditioned systems. *Journal of Process Control*, 18:582–593, 2008. [39]

P. Misra and M. Nikolaou. Input design for model order determination in subspace identification. *AIChE Journal*, 49(8):2124–2132, 2003. [39]

S. Narasimhan and R. Rengaswamy. Multi-objective input signal design for plant friendly identification of process systems. In *American Control Conference*, volume 6, pages 4891–4896, 2004. [22]

B. Ogunnaike, J. Lemaire, M. Morari, and W. Ray. Advanced multivariable control of a pilot-plant distillation column. *AIChE Journal*, 29(4):632–640, 1983. [41]

B.A. Ogunnaike and W.H. Ray. *Process Dynamics, Modeling and Control*, volume 1. Oxford University Press, 1994. [5]

R. Parker, D. Heemstra, F. Doyle, R. Pearson, and B. Ogunnaike. The identification of nonlinear models for process control using tailored "plant-friendly" input sequences. *Journal of Process Control*, 11(2):237–250, 2001. [21]

R. Pintelon and J. Schoukens. *System Identification: A Frequency Domain Approach*. John Wiley & Sons, 2012. [xviii, 15, 17, 18, 19, 20, 25, 26, 27]

J.G. Proakis and D.G. Manolakis. *Digital Signal Processing: Principles, Algorithms, and Applications*. Prentice Hall, 1996. [24]

D. Rees, D.L. Jones, and D.C. Evans. Practical considerations in the design of multisine test signals for system identification. In *9th IEEE Instrumentation and Measurement Technology Conference, IMTC '92*, pages 174–179, 1992. [16]

M. Ren, D. Wu, J. Zhang, and M. Jiang. Minimum entropy-based cascade control for governing hydroelectric turbines. *Entropy*, 16(6):3136–3148, 2014. [21]

D. Rivera and H. Lee. Plant-friendly system identification: a challenge for the process industries. *IFAC Symposium on System Identification (SYSID)*, 2003. [22, 48]

D. Rivera, H. Lee, H. Mittelmann, and M. Braun. High-purity distillation: Using plant-friendly multisine signals to identify a strongly interactive process. *IEEE Control Syst. Mag.*, 27:72–89, 2007. [32, 48]

D. Rivera, H. Lee, H. Mittelmann, and M. Braun. Constrained multisine input signals for plant-friendly identification of chemical process systems. *Journal of Process Control*, 19(4):623–635, 2009. [18, 26, 40, 41, 42]

J. Schoukens, R. Pintelon, and Y. Rolain. *Mastering System Identification in 100 Exercises*. John Wiley & Sons, 2012. [xvi, xviii, 17, 23, 25, 27, 28, 30]

M. Schroeder. Synthesis of low-peak-factor signals and binary sequences with low autocorrelation (corresp.). *IEEE Transactions on Information Theory*, 16(1):85–89, 1970. [20, 25, 27, 45, 48]

S. Skogestad and M. Morari. Lv-control of a high purity distillation column. *Chem. Eng. Sci.*, 43: 33–48, 1988a. [89]

S. Skogestad and M. Morari. Understanding the dynamic behavior of distillation columns. *Ind. Eng. Chem. Res.*, 27(10):1848–1862, 1988b. [89]

S. Skogestad, M. Morari, and J. Doyle. Robust control of ill-conditioned plants: high-purity distillation. *IEEE Transactions on Automatic Control*, 33(12):1092–1105, 1988. [37, 49]

T. Söderström and P. Stoica. *System Identification*. Prentice-Hall, Inc., 1989. [5, 11, 17, 18]

R. Steenis. *Plant-friendly Input Signal Design for System Identification and Robust Control Performance*. PhD thesis, 2009. [xviii, 24]

A. Tan. Direction-dependent systems : A survey. *Automatica*, 45(12):2729–2743, 2009. [40, 41, 42]

A. Van den Bos. Construction of binary multifrequency test signals. In *Proc. 1st IFAC Symp. Identification in Automatic Control Systems*, 1967. [16]

A. Van den Bos. A new method for synthesis of low-peak factor signals. *IEEE Trans. Acoustics, Speech and Signal Processing*, 35(1):120–122, 1987. [16]

P.M.J. Van den Hof, P.S.C. Heuberger, and J. Bokor. System identification with generalized orthonormal basis functions. *Automatica*, 31(12):1821–1834, 1995. [35]

E. Van der Ouderaa, J. Schoukens, and J. Renneboog. Peak factor minimisation of input and output signals of linear systems. *IEEE Trans. Instrumentation and Measurement*, 37(2):207–212, 1988. [22]

V. Vasnani. *Towards Relay Feedback Auto-Tuning of Multi-Loop Systems*. PhD thesis, National University of Singapore, 1994. [90]

R. Wood and M. Berry. Terminal composition control of a binary distillation column. *Chem. Eng. Sci.*, 28:1707–1717, 1973. [41, 89]

Y. Zhu. *Multivariable System Identification for Process Control*. Elsevier, 2001. [40, 41, 42]

Y. Zhu. System identification for process control: Recent experience and outlook. *International Journal of Modelling, Identification and Control*, 6(2):89–103, 2009. [40, 41, 42, 48]

Y. Zhu and P. Stec. Simple control-relevant identification test methods for a class of ill-conditioned processes. *Journal of Process Control*, 16(10):1113–1120, 2006. [30]

# Appendices

# Distillation column simulator

Detail discussion on the parameter tuning can be found in *Paper V*.

## A.1 Distillation column simulator modeling

### A.1.1 General tray

The total material balance for a general tray $i$ is

$$\frac{dM_i}{dt} = V_{i+1} - V_i + Li - 1 - L_i + F_i, \quad i = 1,\ldots,N \tag{A.1.1}$$

**variables:** $M_i$ = is holdups of liquid on tray i [kmol],

$V_{i+1}$ = flow rate of incoming vapor stream [kmol/h]; $V_{N+1}$ = flow rate from reboiler [kmol/h]; $V_i$ = flow rate of outgoing vapor stream [kmol/h];

$L_{i-1}$ = flow rate of ingoing liquid stream [kmol/h]; $L_0$ = flow rate of reflux [kmol/h]; $L_i$ = flow rate of outgoing liquid stream [kmol/h];

$F_i$ = flow rate of feed stream [kmol/h],

At constant pressure, $L_i$ depends on other variables

$$L_i = f(M_i, x_i, V_i) \tag{A.1.2}$$

Typically, in practice, we use Francis weir formula, which is independent of $V_i$. The material balance of the light component (here, ethanol) is

$$M_i \frac{dx_i}{dt} = V_{i+1}y_{i+1} - V_i y_i + Li - 1 x_{i-1} - L_i x_i + F_i z_i, \quad i = 1,\ldots,N \tag{A.1.3}$$

**variables:** $x_i$ = fraction of light component in liquid on tray i;

$y_{i+1}$ = fraction of light component in incoming vapor; $y_{N+1}$ = fraction of light component from reboiler; $y_i$ fraction of light component in outgoing vapor;

$x_{i-1}$ = fraction of light component in incoming liquid; $x_0$ = fraction of light component in reflux;

$z_i$ = fraction of light component in feed stream (applicable only for a feed tray);

The Vapor-liquid equilibriums can be modeled as follows. The fraction of ethanol in the outgoing vapor flow is

$$y_i = E_i y_i^* + (1 - E_i) y_{i+1}, \quad y_i^* = \frac{\alpha_i x_i}{1 + x_i (\alpha_i - 1)} \tag{A.1.4}$$

where $E_i$ is Murphree's tray efficiency (in principle $0 < E_i < 1$ ) and $\alpha_i$ is the relative volatility , which may depend nonlinearly on $x_i$. $E_i$ is tuned based on the column data. The energy balance can be formulate as [Häggblom, 1991]

$$V_i h V_i = V_{i+1} h_{V_{i+1}} + L_{i-1} q_{L_{i-1}} h_{L_{i-1}} + F_i h_{F_i} h_{F_i} - Q_i \tag{A.1.5}$$

where $Q_i$ is the heat loss from the tray through conduction, $h_S$ is the heat of vaporization of a stream $S$ and $q_S$ is the thermal condition of the stream such that $q_S h_S$ is the difference between the specific or molar enthalpy of $S$ and corresponding enthalpy of a saturated liquid with the same composition as $S$ (e.g., $hV_i$ = specific heat of vaporization of outgoing vapor stream [kJ/kmol]). This means that $q < 0$ for a subcooled liquid, $q = 0$ for a saturated liquid, $0 < q < 1$ for a partially vaporised feed, and $q = 1$ for a saturated vapor.

### A.1.2   Condenser

The total material balance of the reboiler is

$$\frac{dM_0}{dt} = V_1 - L - D \tag{A.1.6}$$

The material balance is

$$M_0 \frac{dx_0}{dt} = V_1 y_1 - (L + D) x_D \tag{A.1.7}$$

The energy balance is

$$\frac{dU_0}{dt} = V_1 h_{V_1} - (L + D) \frac{U_0}{M_0} - Q_0. \quad q_L h_L = \frac{U_0}{M_0} \tag{A.1.8}$$

where $M_0$ is the total component holdups and $m_0$ is light component (ethanol) holdups in the reflux drum. For the tray 1, $L_0$ and $x_0$ are equivalent to $L$ and $x_D$, respectively. The heat removal in the condenser can be defined as

$$Q_0 = k_c h_{V_1} V_1 + C \ [kJ/h] \tag{A.1.9}$$

where $k_c$ and $C$ are two adjustable parameters.

### A.1.3   Reboiler

The reboiler is a vertical thermosyphon reboiler, detail can be found in Paper V. According to the construction a fraction ($r$) of the liquid might return to a compartment connected to the reboiler. Assuming constant density of the liquid in the compartment and immediate mixing of the liquid, the balance equations of the reboiler are

$$L_b = L_N - (1 - rR)L_r \tag{A.1.10}$$

where $R = LN + 1/L_r$,

$$x_b = \frac{L_N x_N + rRL_r x_{N+1}}{L_N + rRL_r} \tag{A.1.11}$$

$$M_r \frac{dx_r}{dt} = L_r(x_b - x_r) \tag{A.1.12}$$

The materials balances for the bottom product sump, are

$$\frac{dM_B}{dt} = L_b + (1 - r)RL_r - B \tag{A.1.13}$$

$$\frac{dm_B}{dt} = L_b x_b + (1 - r)RL_r x_{N+1} - Bx_B \tag{A.1.14}$$

where $x_B = \frac{m_B}{M_B}$. For the reboiler, we assume that the dynamic relationship between the steam flow rate $V$ and the heat input $Q_r$ from the condensing steam is first order dynamics with time constant $\tau_r$, and we obtained

$$\tau_r \frac{dQ_r}{dt} = h_V V - Q_r \tag{A.1.15}$$

$$V_{N+1} h_{V_{N+1}} = Q_r \tag{A.1.16}$$

$$L_r = \frac{VN + 1}{1 - R} \tag{A.1.17}$$

$$x_r = Rx_{N+1} + (1 - R)y_{N+1} \tag{A.1.18}$$

$$y_{N+1} = E_r y_{N+1}^* + (1 - E_r)x_r \tag{A.1.19}$$

$$y_{N+1}^* = \frac{\alpha_{N+1} x_{N+1}}{1 + x_{N+1}(\alpha_{N+1} - 1)} \tag{A.1.20}$$

For a given set of inputs ($L$, $V$, $F$, $z_F$, and $T_F$, or ($q_F$)), one can determine the steady state solution $x_D$ and $x_B$ using Eq. A.1.1–A.1.20.

## A.2   Model library

Model library has been created to facilitate different types of model for different purpose of applications as shown in Figure A.1.

The simulator is constructed by combining various trays, reboiler and condenser system, as shown in Appendix A.2. The subsystem models are shown in A.2b. Using the modularity of model library we can easily reuse the tray models for all general trays and their subsystems, however equations are little bit different in feed-tray as shown in fig A.2a. Similarly, the model of condenser and reboiler system have shown in fig A.2b and fig A.2c.



**Figure A.1:** Snapshot of the simulator model library

**Table A.1:** Simulator sub-systems and trays: (a) main interface, (b) trays and subsystems, and (c) model of a general tray.
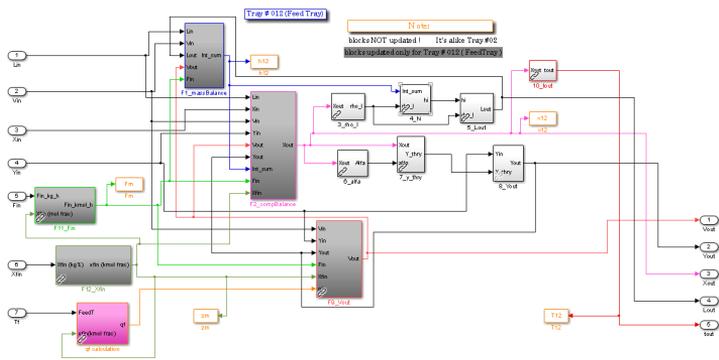
(a) feed tray
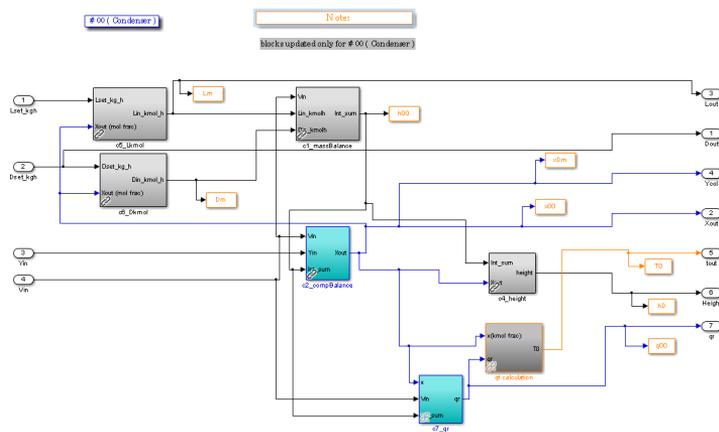


(b) condenser system
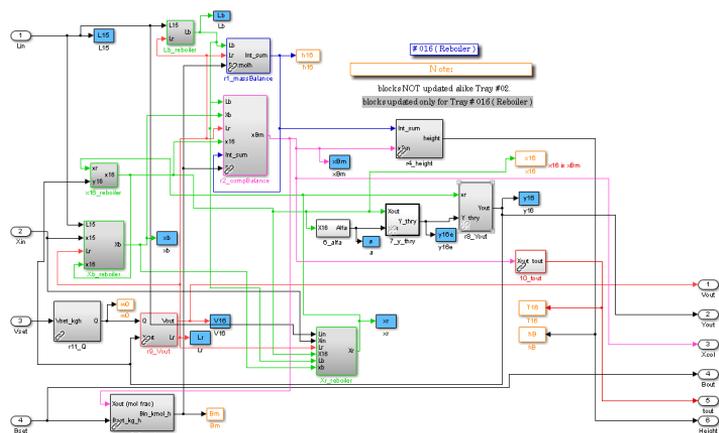


(c) reboiler system

**Figure A.2:** Inside simulator: specific subsystems

## A.3  Parameter tuning with real-life data

Here we have shown in Figures A.3 to A.5, the model performance with few experimental data sets from real life experiments [Häggblom and Böling, 1998]. Experimental data has been compared with the top $y_D$ and bottom concentration $x_B$ and the temperature at topmost tray 1 $T_{01}$. First, 100 data points with sample time $T_s = 30sec$ (30sec*100) are added to acquire the desired steady state before experimental L and V dynamics added into simulator.
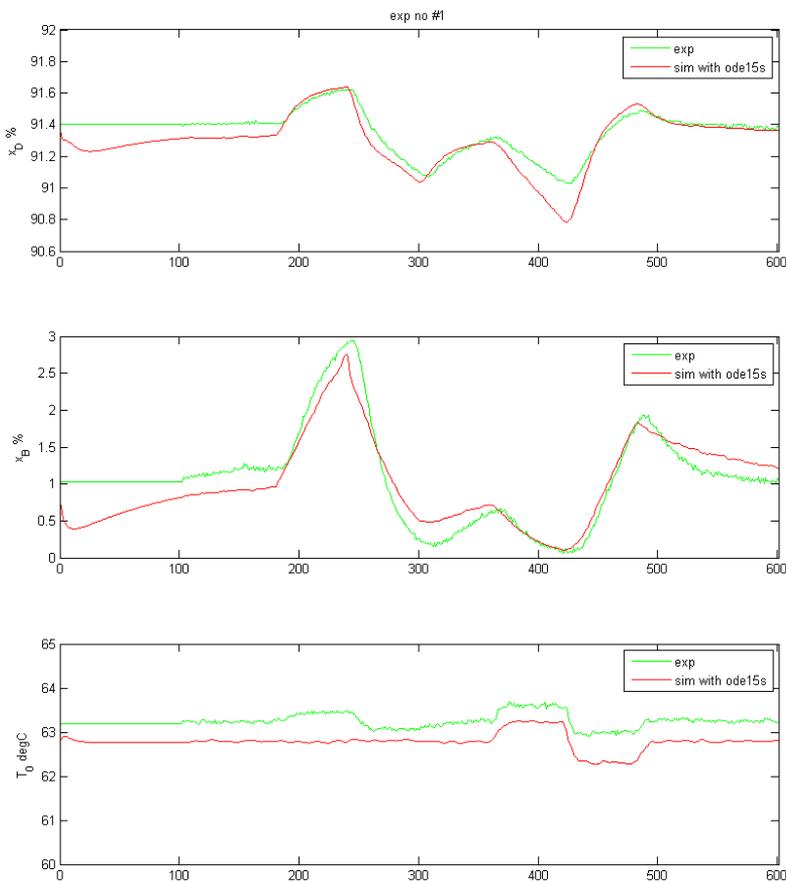


**Figure A.3:** Simulator versus experimental data (Exp. A)
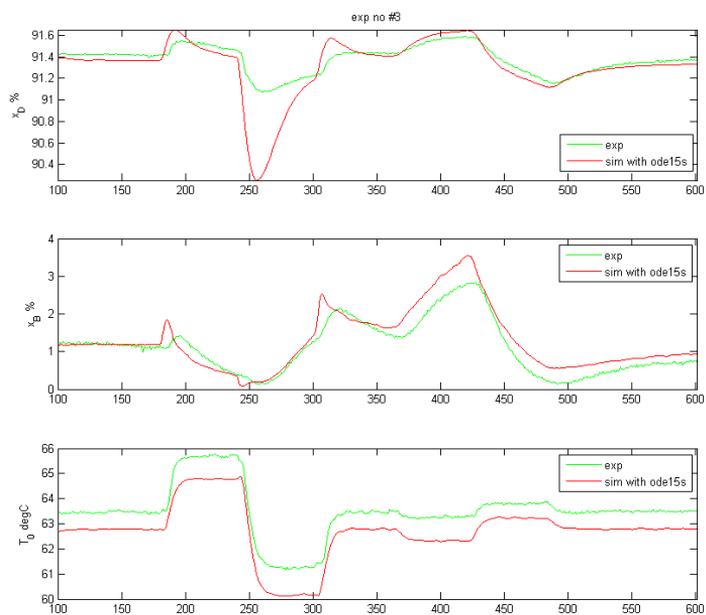
**Figure A.4:** Simulator versus experimental data (Exp. B)
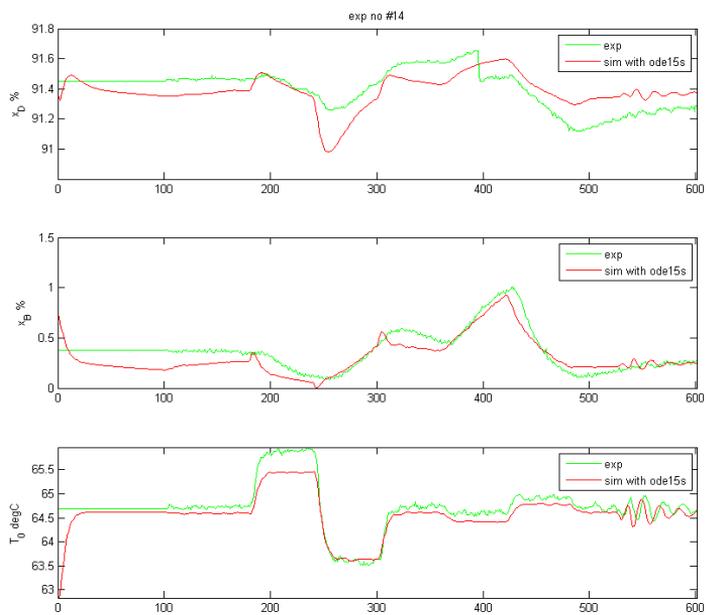


**Figure A.5:** Simulator versus experimental data (Exp. C)

### A.4 Time delay in a state space model

We have formulated a state space system equation with couple of inter variable and have shown with a case study that there is a possibility to achieve better performance than modeling the system with higher order modeling in order to compensate the effect of time delays. Interestingly we can formulate the internal relations from inputs to outputs, among outputs. From their sensitivity analysis, as shown in Figure A.6, we can decide whether a optimization for more accurate time delay is worthy or not. Here each color represents different experimental data.
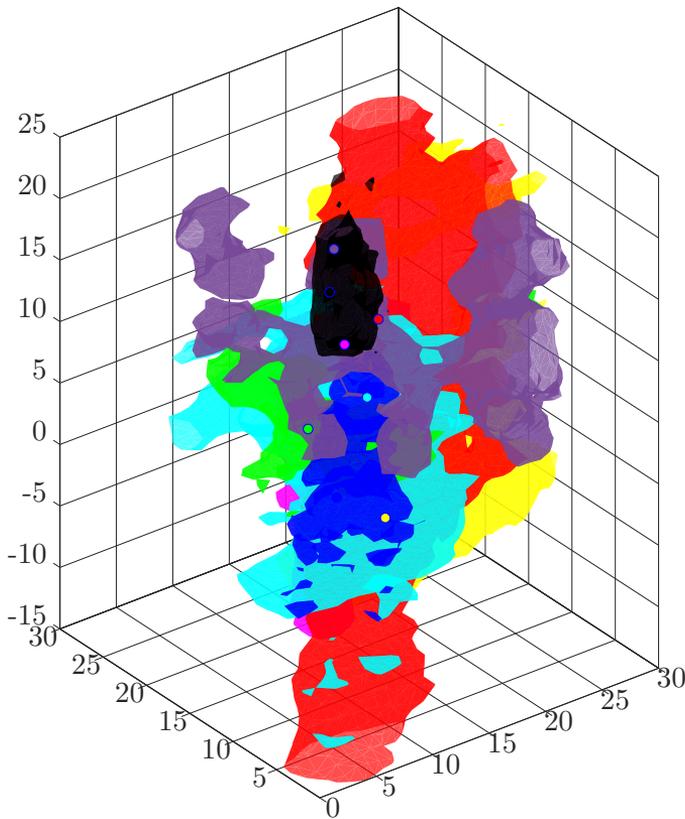


**Figure A.6:** 3D plot for sensitivity of time delays

# Examples of ill-conditioned process models

$T$he case studies (or models) from the literature are summarized here.

### Example 1

A $2 \times 2$ a simplified linear model of a heat exchanger [Jacobsen and Skogestad, 1994].

$$\begin{bmatrix} y_1(s) \\ y_2(s) \end{bmatrix} = \frac{89.243}{(100s+1)(2.439s+1)} \begin{bmatrix} -21(4.76s+1) & 20 \\ & \\ -20 & 21(4.76s+1) \end{bmatrix} \begin{bmatrix} u_1(s) \\ u_2(s) \end{bmatrix} \tag{B.0.1}$$

### Example 2

A $2 \times 2$ reduced $2^{nd}$ order model of a distillation column [Skogestad and Morari, 1988a,b] having condition number $\kappa = 143$.

$$\begin{bmatrix} y_1(s) \\ y_2(s) \end{bmatrix} = \begin{bmatrix} \frac{87.8}{194s+1} & \frac{-87.8}{194s+1} + \frac{1.4}{15s+1} \\ \frac{108.2}{194s+1} & \frac{-108.2}{194s+1} - \frac{1.4}{15s+1} \end{bmatrix} \begin{bmatrix} u_1(s) \\ u_2(s) \end{bmatrix} \tag{B.0.2}$$

### Example 3

A $2 \times 2$ distillation column system, popularly known as *Wood–Berry column* with condition number $\kappa =$ [Wood and Berry, 1973]

$$\begin{bmatrix} y_1(s) \\ y_2(s) \end{bmatrix} = \begin{bmatrix} \dfrac{12.8e^{-1s}}{16.7s+1} & \dfrac{-18.9e^{-3s}}{21s+1} \\[2ex] \dfrac{6.6e^{-7s}}{10.9s+1} & \dfrac{-19.4e^{-3s}}{14.4s+1} \end{bmatrix} \begin{bmatrix} u_1(s) \\ u_2(s) \end{bmatrix} \tag{B.0.3}$$

### Example 4

A $3 \times 3$ model of distillation column [Koung and MacGregor, 1994; Li and Lee, 1996] modified from [Vasnani, 1994] and used in [Häggblom and Ghosh, 2015] with condition number $\kappa = 30$.

$$\begin{bmatrix} y_1(s) \\ y_2(s) \\ y_3(s) \end{bmatrix} = \begin{bmatrix} \dfrac{6e^{-5s}}{22s+1} & \dfrac{20e^{-5s}}{337s+1} & \dfrac{-1e^{-5s}}{10s+1} \\[2ex] \dfrac{8e^{-5s}}{50s+1} & \dfrac{77e^{-3s}}{28s+1} & \dfrac{-5e^{-5s}}{10s+1} \\[2ex] \dfrac{9e^{-5s}}{50s+1} & \dfrac{-37e^{-5s}}{166s+1} & \dfrac{-103e^{-4s}}{23s+1} \end{bmatrix} \begin{bmatrix} u_1(s) \\ u_2(s) \\ u_3(s) \end{bmatrix} \tag{B.0.4}$$

### Example 5

A $4 \times 4$ model of distillation column system* [Alatiqi and Luyben, 1986; Ghosh, 2014; Ghosh et al., 2014]. The system has high condition number $\kappa = 125$.

$$G_5(s) = \begin{bmatrix} \dfrac{4.09e^{-1.3s}}{(33s+1)(8.3s+1)} & \dfrac{6.36e^{-1.2s}}{(31.6s+1)(20s+1)} & \dfrac{-0.25e^{-1.4s}}{21s+1} & \dfrac{-0.49e^{-6s}}{22s+1} \\[2ex] \dfrac{-4.17e^{-5s}}{45s+1} & \dfrac{6.93e^{-1.02s}}{44.6s+1} & \dfrac{-0.05e^{-6s}}{(34.5s+1)^2} & \dfrac{1.53e^{-3.8s}}{48s+1} \\[2ex] \dfrac{1.73e^{-18s}}{(13s+1)^2} & \dfrac{5.11e^{-12s}}{(13.3s+1)^2} & \dfrac{4.61e^{-1.01s}}{18.5s+1} & \dfrac{-5.49e^{-1.5s}}{15s+1} \\[2ex] \dfrac{-11.2e^{-2.6s}}{(43s+1)(6.5s+1)} & \dfrac{14(10s+1)e^{-0.02s}}{(45s+1)(17.4s^2+3s+1)} & \dfrac{-0.1e^{-0.05s}}{(31.6s+1)(5s+1)} & \dfrac{4.49e^{-0.6s}}{(48s+1)(6.3s+1)} \end{bmatrix} \tag{B.0.5}$$

### Example 6

Here is an interesting example where the system is ill-conditioned not only because of its smallest $\sigma(\sigma_4)$, but also for $\sigma_3$. Interestingly both $\sigma$ s $(\sigma_3, \sigma_4)$ are atleast 110 times smaller than $\sigma_1$. The singular values of this system are $\Sigma = \text{diag}\{20.39, 16.71, 0.185, 0.156\}$.

$$G_6(s) = \begin{bmatrix} \dfrac{0.173e^{-3.79s}}{(21.74s+1)^2} & \dfrac{-3.4e^{-7.75s}}{(22.22s+1)^2} & \dfrac{-8.66e^{-1.59s}}{11.36s+1} & \dfrac{11.94e^{-27.33s}}{33.3s+1} \\[2ex] \dfrac{-0.268e^{-60s}}{400s+1} & \dfrac{2.79e^{-0.71s}}{(66.67s+1)^2} & \dfrac{5.05e^{-2.24s}}{14.29s+1} & \dfrac{-5.65e^{-8.72s}}{(250s+1)^2} \\[2ex] \dfrac{-0.084e^{-0.68s}}{(2.38s+1)^2} & \dfrac{0.94e^{-0.59s}}{(7.14s+1)^2} & \dfrac{1.31e^{-0.42s}}{(1.43s+1)^2} & -0.56e^{-s} \\[2ex] \dfrac{-0.047e^{-0.52s}}{11.11s+1} & \dfrac{-15.5e^{-0.48s}}{(6.90s+1)^2} & \dfrac{-9.03e^{-1.91s}}{12.19s+1} & -8.71e^{-s} \end{bmatrix} \tag{B.0.6}$$

---

*More detail discussion on the system can be found in *Paper II–IV* and in Chapter 3

# Estimated models of $4 \times 4$ case-study

Estimated transfer function matrix models from the case study 2 are listed here (*two pole with time delay model*, **2P+TD**), as in Equation (6.1.1)

$$G_{ij} = \frac{K_p \ e^{-sT_d}}{(1 + sT_{p_1})(1 + sT_{p_2})}$$

where $G_{ij}$ is the $i^{th}$ row and $j^{th}$ column element of the transfer function matrix of the system (as shown in Equation (6.3.1)). For detail description of experiments and their design methods please refer to Chapter 6, particularly at Table 6.1.

**Table C.1:** Estimated model of $4 \times 4$ case study

| Exp | Esimated model | | | |
|---|---|---|---|---|
| 1 | $\dfrac{4.10e^{-2.12s}}{(32.75s+1)(7.69s+1)}$ $\dfrac{-4.15e^{-3.87s}}{(45.18s+1)(0.49s+1)}$ $\dfrac{1.73e^{-15.10s}}{(14.65s+1)(14.65s+1)}$ $\dfrac{-11.20e^{-2.39s}}{(42.82s+1)(6.77s+1)}$ | $\dfrac{6.36e^{-0.42s}}{(26.02s+1)(26.02s+1)}$ $\dfrac{6.96e^{-0.00s}}{(44.43s+1)(1.20s+1)}$ $\dfrac{5.15e^{-14.32s}}{(17.11s+1)(7.05s+1)}$ $\dfrac{13.75e^{-0.00s}}{(0.00s+1)(36.06s+1)}$ | $\dfrac{-0.25e^{-1.56s}}{(19.35s+1)(0.06s+1)}$ $\dfrac{-0.05e^{-16.94s}}{(22.53s+1)(22.53s+1)}$ $\dfrac{4.61e^{-0.00s}}{(18.42s+1)(1.16s+1)}$ $\dfrac{-0.10e^{-1.59s}}{(22.94s+1)(9.99s+1)}$ | $\dfrac{-0.49e^{-5.62s}}{(21.91s+1)(0.32s+1)}$ $\dfrac{1.53e^{-3.82s}}{(48.24s+1)(0.04s+1)}$ $\dfrac{-5.49e^{-0.18s}}{(14.42s+1)(1.60s+1)}$ $\dfrac{4.49e^{-0.73s}}{(48.03s+1)(6.10s+1)}$ |
| 2 | $\dfrac{4.08e^{-1.69s}}{(33.27s+1)(7.49s+1)}$ $\dfrac{-4.17e^{-4.90s}}{(45.03s+1)(0.50s+1)}$ $\dfrac{1.72e^{-16.69s}}{(16.69s+1)(10.11s+1)}$ $\dfrac{-11.22e^{-2.66s}}{(42.52s+1)(7.11s+1)}$ | $\dfrac{6.35e^{-0.45s}}{(26.26s+1)(26.26s+1)}$ $\dfrac{6.91e^{-0.40s}}{(44.59s+1)(0.89s+1)}$ $\dfrac{5.11e^{-14.48s}}{(17.60s+1)(6.63s+1)}$ $\dfrac{13.78e^{-0.64s}}{(0.00s+1)(36.03s+1)}$ | $\dfrac{-0.26e^{-2.04s}}{(20.94s+1)(0.04s+1)}$ $\dfrac{-0.05e^{-12.87s}}{(38.33s+1)(38.33s+1)}$ $\dfrac{4.61e^{-0.00s}}{(18.40s+1)(0.61s+1)}$ $\dfrac{-0.12e^{-2.16s}}{(34.26s+1)(34.59s+1)}$ | $\dfrac{-0.48e^{-5.78s}}{(20.72s+1)(0.34s+1)}$ $\dfrac{1.53e^{-4.11s}}{(48.58s+1)(0.05s+1)}$ $\dfrac{-5.48e^{-0.00s}}{(14.69s+1)(1.25s+1)}$ $\dfrac{4.50e^{-2.48s}}{(46.53s+1)(6.73s+1)}$ |
| 3 | $\dfrac{4.11e^{-1.65s}}{(33.92s+1)(7.50s+1)}$ $\dfrac{-4.20e^{-4.94s}}{(45.37s+1)(0.02s+1)}$ $\dfrac{1.74e^{-18.06s}}{(16.62s+1)(9.32s+1)}$ $\dfrac{-11.27e^{-2.94s}}{(45.11s+1)(5.59s+1)}$ | $\dfrac{6.34e^{-0.97s}}{(25.59s+1)(25.59s+1)}$ $\dfrac{6.96e^{-0.94s}}{(45.00s+1)(0.08s+1)}$ $\dfrac{5.11e^{-12.17s}}{(13.15s+1)(13.33s+1)}$ $\dfrac{12.39e^{-0.00s}}{(0.00s+1)(23.23s+1)}$ | $\dfrac{-0.25e^{-1.52s}}{(21.14s+1)(0.04s+1)}$ $\dfrac{-0.05e^{-9.10s}}{(37.50s+1)(37.50s+1)}$ $\dfrac{4.61e^{-0.93s}}{(18.47s+1)(0.12s+1)}$ $\dfrac{-0.25e^{-0.02s}}{(42.95s+1)(1.66s+1)}$ | $\dfrac{-0.48e^{-5.97s}}{(22.10s+1)(0.02s+1)}$ $\dfrac{1.52e^{-3.82s}}{(47.18s+1)(0.00s+1)}$ $\dfrac{-5.47e^{-1.10s}}{(14.98s+1)(0.45s+1)}$ $\dfrac{4.44e^{-0.74s}}{(47.63s+1)(6.26s+1)}$ |
| 4 | $\dfrac{4.10e^{-1.60s}}{(33.64s+1)(7.63s+1)}$ $\dfrac{-4.17e^{-4.97s}}{(45.10s+1)(0.02s+1)}$ $\dfrac{1.72e^{-18.17s}}{(16.71s+1)(9.63s+1)}$ $\dfrac{-11.11e^{-2.67s}}{(42.88s+1)(6.52s+1)}$ | $\dfrac{6.34e^{-0.97s}}{(25.50s+1)(25.50s+1)}$ $\dfrac{6.93e^{-0.94s}}{(44.53s+1)(0.08s+1)}$ $\dfrac{5.16e^{-12.03s}}{(11.76s+1)(14.79s+1)}$ $\dfrac{11.98e^{-0.00s}}{(0.00s+1)(22.63s+1)}$ | $\dfrac{-0.25e^{-1.52s}}{(20.47s+1)(0.01s+1)}$ $\dfrac{-0.06e^{-1.91s}}{(41.91s+1)(41.91s+1)}$ $\dfrac{4.62e^{-1.00s}}{(18.50s+1)(0.00s+1)}$ $\dfrac{-0.15e^{-0.20s}}{(50.00s+1)(1.02s+1)}$ | $\dfrac{-0.48e^{-5.98s}}{(21.97s+1)(0.02s+1)}$ $\dfrac{1.54e^{-3.78s}}{(48.06s+1)(0.00s+1)}$ $\dfrac{-5.49e^{-1.38s}}{(14.90s+1)(0.15s+1)}$ $\dfrac{4.43e^{-0.79s}}{(48.07s+1)(6.13s+1)}$ |
| 5 | $\dfrac{4.13e^{-1.26s}}{(33.01s+1)(8.10s+1)}$ $\dfrac{-4.21e^{-4.76s}}{(45.50s+1)(0.25s+1)}$ $\dfrac{1.81e^{-17.86s}}{(15.18s+1)(11.57s+1)}$ $\dfrac{-10.19e^{-1.81s}}{(25.71s+1)(11.88s+1)}$ | $\dfrac{6.35e^{-1.56s}}{(18.59s+1)(33.45s+1)}$ $\dfrac{7.05e^{-0.50s}}{(45.44s+1)(0.49s+1)}$ $\dfrac{5.08e^{-11.94s}}{(11.36s+1)(15.25s+1)}$ $\dfrac{12.17e^{-0.98s}}{(0.00s+1)(21.94s+1)}$ | $\dfrac{-0.24e^{-1.37s}}{(19.74s+1)(0.00s+1)}$ $\dfrac{-0.04e^{-9.79s}}{(34.49s+1)(34.49s+1)}$ $\dfrac{4.55e^{-1.05s}}{(18.13s+1)(0.00s+1)}$ $\dfrac{-0.28e^{-4.35s}}{(2.53s+1)(2.52s+1)}$ | $\dfrac{-0.51e^{-5.97s}}{(22.45s+1)(0.00s+1)}$ $\dfrac{1.54e^{-3.78s}}{(48.58s+1)(0.00s+1)}$ $\dfrac{-5.45e^{-1.44s}}{(14.83s+1)(0.10s+1)}$ $\dfrac{4.24e^{-0.37s}}{(32.17s+1)(6.35s+1)}$ |
| 6 | $\dfrac{4.05e^{-1.36s}}{(33.01s+1)(8.24s+1)}$ $\dfrac{-4.17e^{-4.99s}}{(45.23s+1)(0.02s+1)}$ $\dfrac{1.75e^{-17.94s}}{(13.15s+1)(13.21s+1)}$ $\dfrac{-9.96e^{-2.67s}}{(25.76s+1)(11.65s+1)}$ | $\dfrac{6.37e^{-0.83s}}{(25.67s+1)(25.67s+1)}$ $\dfrac{6.96e^{-0.95s}}{(44.74s+1)(0.08s+1)}$ $\dfrac{5.01e^{-12.11s}}{(11.91s+1)(14.79s+1)}$ $\dfrac{11.36e^{-1.64s}}{(0.00s+1)(20.21s+1)}$ | $\dfrac{-0.25e^{-1.48s}}{(21.50s+1)(0.01s+1)}$ $\dfrac{-0.05e^{-1.91s}}{(33.71s+1)(35.50s+1)}$ $\dfrac{4.57e^{-1.06s}}{(18.37s+1)(0.00s+1)}$ $\dfrac{-0.25e^{-5.09s}}{(2.38s+1)(2.39s+1)}$ | $\dfrac{-0.47e^{-6.01s}}{(21.20s+1)(0.02s+1)}$ $\dfrac{1.53e^{-3.80s}}{(47.98s+1)(0.00s+1)}$ $\dfrac{-5.46e^{-1.38s}}{(14.93s+1)(0.17s+1)}$ $\dfrac{4.06e^{-1.39s}}{(31.98s+1)(6.20s+1)}$ |

| Exp | Esimated model | | | |
|-----|---|---|---|---|
| 7 | $\dfrac{4.10e^{-1.32s}}{(33.11s+1)(8.23s+1)}$ $\dfrac{-4.19e^{-4.97s}}{(45.40s+1)(0.01s+1)}$ $\dfrac{1.75e^{-18.14s}}{(16.17s+1)(10.20s+1)}$ $\dfrac{-10.46e^{-0.61s}}{(32.26s+1)(12.78s+1)}$ | $\dfrac{6.41e^{-1.09s}}{(19.63s+1)(32.28s+1)}$ $\dfrac{6.91e^{-1.00s}}{(44.07s+1)(0.03s+1)}$ $\dfrac{5.21e^{-12.24s}}{(9.78s+1)(17.02s+1)}$ $\dfrac{10.23e^{-1.40s}}{(0.00s+1)(16.22s+1)}$ | $\dfrac{-0.25e^{-1.45s}}{(21.03s+1)(0.01s+1)}$ $\dfrac{-0.05e^{-6.97s}}{(47.11s+1)(26.39s+1)}$ $\dfrac{4.59e^{-1.01s}}{(18.42s+1)(0.00s+1)}$ $\dfrac{-0.17e^{-5.88s}}{(1.80s+1)(1.81s+1)}$ | $\dfrac{-0.49e^{-6.01s}}{(21.94s+1)(0.02s+1)}$ $\dfrac{1.54e^{-3.77s}}{(48.42s+1)(0.00s+1)}$ $\dfrac{-5.47e^{-1.47s}}{(14.93s+1)(0.04s+1)}$ $\dfrac{4.08e^{-0.02s}}{(39.14s+1)(6.23s+1)}$ |
| 8 | $\dfrac{4.08e^{-1.33s}}{(32.89s+1)(8.42s+1)}$ $\dfrac{-4.19e^{-5.00s}}{(45.16s+1)(0.01s+1)}$ $\dfrac{1.74e^{-18.09s}}{(14.24s+1)(11.64s+1)}$ $\dfrac{-10.90e^{-1.40s}}{(37.72s+1)(11.22s+1)}$ | $\dfrac{6.36e^{-1.05s}}{(21.57s+1)(29.72s+1)}$ $\dfrac{6.91e^{-1.00s}}{(44.88s+1)(0.00s+1)}$ $\dfrac{5.09e^{-12.28s}}{(12.97s+1)(13.17s+1)}$ $\dfrac{11.19e^{-0.95s}}{(0.00s+1)(19.20s+1)}$ | $\dfrac{-0.24e^{-1.40s}}{(20.94s+1)(0.01s+1)}$ $\dfrac{-0.05e^{-6.05s}}{(34.06s+1)(37.13s+1)}$ $\dfrac{4.61e^{-1.02s}}{(18.48s+1)(0.00s+1)}$ $\dfrac{-0.17e^{-5.54s}}{(2.09s+1)(1.58s+1)}$ | $\dfrac{-0.49e^{-6.01s}}{(21.99s+1)(0.02s+1)}$ $\dfrac{1.53e^{-3.80s}}{(48.18s+1)(0.00s+1)}$ $\dfrac{-5.49e^{-1.42s}}{(14.99s+1)(0.09s+1)}$ $\dfrac{4.28e^{-0.00s}}{(42.47s+1)(6.25s+1)}$ |
| 9 | $\dfrac{4.10e^{-1.60s}}{(33.84s+1)(7.66s+1)}$ $\dfrac{-4.13e^{-4.96s}}{(44.36s+1)(0.01s+1)}$ $\dfrac{1.72e^{-17.67s}}{(13.12s+1)(13.06s+1)}$ $\dfrac{-11.08e^{-2.56s}}{(42.37s+1)(6.64s+1)}$ | $\dfrac{6.34e^{-0.97s}}{(25.51s+1)(25.51s+1)}$ $\dfrac{6.93e^{-1.00s}}{(44.73s+1)(0.06s+1)}$ $\dfrac{5.17e^{-12.02s}}{(11.24s+1)(15.25s+1)}$ $\dfrac{11.91e^{-0.48s}}{(0.00s+1)(24.28s+1)}$ | $\dfrac{-0.24e^{-1.52s}}{(19.94s+1)(0.00s+1)}$ $\dfrac{-0.04e^{-2.34s}}{(35.73s+1)(35.73s+1)}$ $\dfrac{4.61e^{-1.00s}}{(18.58s+1)(0.00s+1)}$ $\dfrac{-0.08e^{-5.19s}}{(25.40s+1)(0.76s+1)}$ | $\dfrac{-0.48e^{-5.99s}}{(21.83s+1)(0.02s+1)}$ $\dfrac{1.53e^{-3.79s}}{(47.89s+1)(0.00s+1)}$ $\dfrac{-5.47e^{-1.32s}}{(14.83s+1)(0.21s+1)}$ $\dfrac{4.44e^{-0.59s}}{(46.93s+1)(6.41s+1)}$ |
| 10 | $\dfrac{3.99e^{-1.48s}}{(32.74s+1)(7.96s+1)}$ $\dfrac{-4.21e^{-4.98s}}{(44.84s+1)(0.00s+1)}$ $\dfrac{1.75e^{-18.20s}}{(14.74s+1)(11.58s+1)}$ $\dfrac{-10.30e^{-1.49s}}{(27.40s+1)(12.36s+1)}$ | $\dfrac{6.39e^{-1.16s}}{(25.33s+1)(25.33s+1)}$ $\dfrac{6.91e^{-0.96s}}{(43.97s+1)(0.08s+1)}$ $\dfrac{5.16e^{-11.98s}}{(13.32s+1)(13.26s+1)}$ $\dfrac{11.92e^{-0.69s}}{(0.00s+1)(24.28s+1)}$ | $\dfrac{-0.24e^{-1.44s}}{(20.84s+1)(0.01s+1)}$ $\dfrac{-0.07e^{-2.59s}}{(35.85s+1)(35.85s+1)}$ $\dfrac{4.59e^{-1.01s}}{(18.22s+1)(0.00s+1)}$ $\dfrac{-0.28e^{-3.17s}}{(3.47s+1)(3.46s+1)}$ | $\dfrac{-0.47e^{-6.01s}}{(20.41s+1)(0.02s+1)}$ $\dfrac{1.57e^{-3.81s}}{(48.13s+1)(0.00s+1)}$ $\dfrac{-5.47e^{-1.41s}}{(14.79s+1)(0.10s+1)}$ $\dfrac{4.25e^{-0.00s}}{(33.72s+1)(7.14s+1)}$ |
| 11 | $\dfrac{4.12e^{-1.59s}}{(33.79s+1)(7.71s+1)}$ $\dfrac{-4.12e^{-5.13s}}{(44.15s+1)(0.02s+1)}$ $\dfrac{1.75e^{-17.75s}}{(13.71s+1)(12.95s+1)}$ $\dfrac{-11.07e^{-2.65s}}{(42.33s+1)(6.60s+1)}$ | $\dfrac{6.35e^{-0.96s}}{(25.39s+1)(25.39s+1)}$ $\dfrac{6.92e^{-0.96s}}{(44.47s+1)(0.08s+1)}$ $\dfrac{5.13e^{-12.11s}}{(11.25s+1)(15.75s+1)}$ $\dfrac{12.18e^{-0.00s}}{(0.00s+1)(25.27s+1)}$ | $\dfrac{-0.24e^{-1.52s}}{(20.10s+1)(0.00s+1)}$ $\dfrac{-0.05e^{-1.91s}}{(35.04s+1)(35.04s+1)}$ $\dfrac{4.61e^{-1.00s}}{(18.45s+1)(0.00s+1)}$ $\dfrac{-0.15e^{-4.65s}}{(43.46s+1)(1.77s+1)}$ | $\dfrac{-0.48e^{-5.99s}}{(21.04s+1)(0.01s+1)}$ $\dfrac{1.51e^{-3.84s}}{(47.66s+1)(0.00s+1)}$ $\dfrac{-5.48e^{-1.30s}}{(14.99s+1)(0.21s+1)}$ $\dfrac{4.47e^{-0.68s}}{(47.78s+1)(6.23s+1)}$ |
| 12 | $\dfrac{4.08e^{-1.54s}}{(33.24s+1)(7.91s+1)}$ $\dfrac{-4.22e^{-4.96s}}{(45.46s+1)(0.02s+1)}$ $\dfrac{1.70e^{-18.29s}}{(14.71s+1)(10.72s+1)}$ $\dfrac{-10.12e^{-2.65s}}{(34.69s+1)(7.28s+1)}$ | $\dfrac{6.46e^{-0.85s}}{(25.51s+1)(25.51s+1)}$ $\dfrac{6.95e^{-0.91s}}{(44.59s+1)(0.08s+1)}$ $\dfrac{5.13e^{-12.11s}}{(12.73s+1)(13.44s+1)}$ $\dfrac{12.15e^{-0.47s}}{(0.00s+1)(23.75s+1)}$ | $\dfrac{-0.23e^{-1.47s}}{(20.50s+1)(0.01s+1)}$ $\dfrac{-0.05e^{-1.91s}}{(32.28s+1)(32.28s+1)}$ $\dfrac{4.61e^{-1.08s}}{(18.42s+1)(0.00s+1)}$ $\dfrac{0.24e^{-12.78s}}{(9.24s+1)(3.00s+1)}$ | $\dfrac{-0.50e^{-6.01s}}{(21.47s+1)(0.02s+1)}$ $\dfrac{1.55e^{-3.78s}}{(48.03s+1)(0.00s+1)}$ $\dfrac{-5.48e^{-1.45s}}{(14.92s+1)(0.11s+1)}$ $\dfrac{3.62e^{-0.00s}}{(39.57s+1)(8.51s+1)}$ |
| 13 | $\dfrac{4.10e^{-1.60s}}{(33.91s+1)(7.65s+1)}$ $\dfrac{-4.14e^{-4.98s}}{(44.66s+1)(0.07s+1)}$ $\dfrac{1.73e^{-18.05s}}{(15.41s+1)(11.03s+1)}$ $\dfrac{-11.24e^{-2.55s}}{(43.11s+1)(6.56s+1)}$ | $\dfrac{6.34e^{-0.97s}}{(25.60s+1)(25.60s+1)}$ $\dfrac{6.95e^{-0.87s}}{(44.58s+1)(0.14s+1)}$ $\dfrac{5.16e^{-11.81s}}{(13.85s+1)(13.02s+1)}$ $\dfrac{12.02e^{-0.46s}}{(0.00s+1)(24.57s+1)}$ | $\dfrac{-0.24e^{-1.52s}}{(20.39s+1)(0.01s+1)}$ $\dfrac{-0.04e^{-8.13s}}{(33.16s+1)(33.16s+1)}$ $\dfrac{4.60e^{-1.01s}}{(18.39s+1)(0.00s+1)}$ $\dfrac{-0.10e^{-0.86s}}{(31.02s+1)(4.00s+1)}$ | $\dfrac{-0.48e^{-5.98s}}{(21.81s+1)(0.02s+1)}$ $\dfrac{1.53e^{-3.82s}}{(48.03s+1)(0.00s+1)}$ $\dfrac{-5.50e^{-1.30s}}{(15.02s+1)(0.22s+1)}$ $\dfrac{4.48e^{-0.66s}}{(47.61s+1)(6.37s+1)}$ |

| Exp | Esimated model | | | |
|---|---|---|---|---|
| 14 | $\dfrac{4.10e^{-1.31s}}{(33.08s+1)(8.36s+1)}$ | $\dfrac{6.39e^{-1.58s}}{(18.88s+1)(32.98s+1)}$ | $\dfrac{-0.25e^{-1.45s}}{(20.55s+1)(0.00s+1)}$ | $\dfrac{-0.49e^{-6.01s}}{(22.33s+1)(0.02s+1)}$ |
| | $\dfrac{-4.15e^{-4.96s}}{(44.00s+1)(0.03s+1)}$ | $\dfrac{6.91e^{-1.00s}}{(44.23s+1)(0.00s+1)}$ | $\dfrac{-0.07e^{-1.42s}}{(36.35s+1)(36.35s+1)}$ | $\dfrac{1.55e^{-3.79s}}{(47.42s+1)(0.00s+1)}$ |
| | $\dfrac{1.69e^{-18.06s}}{(13.73s+1)(12.24s+1)}$ | $\dfrac{5.16e^{-11.85s}}{(11.80s+1)(14.86s+1)}$ | $\dfrac{4.66e^{-1.00s}}{(18.57s+1)(0.00s+1)}$ | $\dfrac{-5.53e^{-1.36s}}{(15.03s+1)(0.14s+1)}$ |
| | $\dfrac{-10.71e^{-1.57s}}{(30.18s+1)(11.22s+1)}$ | $\dfrac{12.20e^{-0.82s}}{(0.00s+1)(25.76s+1)}$ | $\dfrac{-0.23e^{-3.86s}}{(3.06s+1)(2.91s+1)}$ | $\dfrac{4.33e^{-0.40s}}{(35.62s+1)(6.98s+1)}$ |
| 15 | $\dfrac{4.09e^{-1.34s}}{(33.01s+1)(8.21s+1)}$ | $\dfrac{6.36e^{-1.13s}}{(20.79s+1)(30.89s+1)}$ | $\dfrac{-0.25e^{-1.24s}}{(20.39s+1)(0.00s+1)}$ | $\dfrac{-0.50e^{-5.79s}}{(22.69s+1)(0.02s+1)}$ |
| | $\dfrac{-4.17e^{-5.10s}}{(44.60s+1)(0.02s+1)}$ | $\dfrac{6.94e^{-1.03s}}{(44.39s+1)(0.08s+1)}$ | $\dfrac{-0.06e^{-1.91s}}{(37.00s+1)(37.00s+1)}$ | $\dfrac{1.53e^{-3.99s}}{(48.20s+1)(0.00s+1)}$ |
| | $\dfrac{1.77e^{-17.51s}}{(13.33s+1)(13.51s+1)}$ | $\dfrac{5.15e^{-12.12s}}{(11.97s+1)(14.49s+1)}$ | $\dfrac{4.59e^{-1.07s}}{(18.26s+1)(0.00s+1)}$ | $\dfrac{-5.50e^{-1.34s}}{(15.04s+1)(0.16s+1)}$ |
| | $\dfrac{-11.10e^{-2.56s}}{(42.19s+1)(6.76s+1)}$ | $\dfrac{12.22e^{-0.00s}}{(0.00s+1)(25.76s+1)}$ | $\dfrac{-0.10e^{-4.75s}}{(29.78s+1)(1.98s+1)}$ | $\dfrac{4.47e^{-0.69s}}{(48.05s+1)(6.15s+1)}$ |
| 16 | $\dfrac{4.07e^{-1.38s}}{(32.91s+1)(8.31s+1)}$ | $\dfrac{6.33e^{-1.02s}}{(25.68s+1)(25.68s+1)}$ | $\dfrac{-0.24e^{-1.42s}}{(21.01s+1)(0.01s+1)}$ | $\dfrac{-0.49e^{-6.05s}}{(21.53s+1)(0.02s+1)}$ |
| | $\dfrac{-4.14e^{-4.97s}}{(44.71s+1)(0.00s+1)}$ | $\dfrac{6.90e^{-1.00s}}{(44.34s+1)(0.04s+1)}$ | $\dfrac{-0.05e^{-8.35s}}{(34.59s+1)(34.90s+1)}$ | $\dfrac{1.52e^{-3.79s}}{(47.76s+1)(0.00s+1)}$ |
| | $\dfrac{1.74e^{-17.75s}}{(13.31s+1)(13.36s+1)}$ | $\dfrac{5.23e^{-11.38s}}{(12.15s+1)(15.06s+1)}$ | $\dfrac{4.61e^{-1.03s}}{(18.42s+1)(0.00s+1)}$ | $\dfrac{-5.48e^{-1.42s}}{(14.92s+1)(0.11s+1)}$ |
| | $\dfrac{-10.37e^{-2.58s}}{(35.34s+1)(7.40s+1)}$ | $\dfrac{12.50e^{-0.49s}}{(0.00s+1)(24.37s+1)}$ | $\dfrac{0.30e^{-13.69s}}{(13.31s+1)(0.07s+1)}$ | $\dfrac{3.66e^{-0.00s}}{(40.50s+1)(8.50s+1)}$ |
| 17 | $\dfrac{4.15e^{-1.38s}}{(33.38s+1)(8.31s+1)}$ | $\dfrac{6.48e^{-1.46s}}{(18.40s+1)(34.36s+1)}$ | $\dfrac{-0.25e^{-1.43s}}{(20.98s+1)(0.00s+1)}$ | $\dfrac{-0.50e^{-6.01s}}{(22.73s+1)(0.01s+1)}$ |
| | $\dfrac{-4.19e^{-4.99s}}{(45.26s+1)(0.02s+1)}$ | $\dfrac{6.89e^{-0.95s}}{(44.05s+1)(0.07s+1)}$ | $\dfrac{-0.05e^{-1.82s}}{(34.50s+1)(35.33s+1)}$ | $\dfrac{1.53e^{-3.78s}}{(47.91s+1)(0.00s+1)}$ |
| | $\dfrac{1.67e^{-18.09s}}{(15.23s+1)(10.09s+1)}$ | $\dfrac{5.12e^{-11.95s}}{(10.58s+1)(14.53s+1)}$ | $\dfrac{4.68e^{-1.00s}}{(18.50s+1)(0.00s+1)}$ | $\dfrac{-5.55e^{-1.40s}}{(14.97s+1)(0.09s+1)}$ |
| | $\dfrac{-11.05e^{-0.76s}}{(37.55s+1)(10.11s+1)}$ | $\dfrac{12.48e^{-0.53s}}{(0.00s+1)(26.12s+1)}$ | $\dfrac{-0.27e^{-2.95s}}{(3.34s+1)(3.33s+1)}$ | $\dfrac{4.52e^{-0.00s}}{(41.94s+1)(5.29s+1)}$ |

# Symbols and Abbreviations

| | |
|---|---|
| $y_i^\times$ | projection of output $Y$ along the $i$–th gain-direction, [43] |
| $\alpha$ | closed-loop response parameter, [18] |
| $\beta$ | settling-time parameter, [18] |
| $\delta$ | notch frequency, [30] |
| $\Phi$ | plant friendliness of an input signal, [21] |
| $\sigma_i$ | $i$-th singular value, [31] |
| $\Upsilon$ | determinant of correlation matrix of system outputs, [vii] |
| $\xi_i$ | ratio of singular values ($\sigma_i$) of a system, where $\xi_i = \sigma_1/\sigma_i$ , [65] |
| $G(s)$ | the transfer function matrix of a system, [38] |
| $N_s$ | sequence length of a signal, [29] |
| $T$ | sampling time, [29] |
| $T_L$ | the experiment length, [55] |
| $T_{sw}$ | switching time of a PRBS, [56] |
| $\Phi_u(j)$ | signal spectrum in frequency domain, [31] |
| $u_j(k)$ | signal spectrum in time domain, [31] |
| $\hat{\delta}_{ji}$ and $\hat{a}_{ji}$ | *snow effect* coefficients, [31] |
| $\tau_H$ | high dominating time constant, [32] |
| $\tau_L$ | low dominating time constant, [32] |
| $\epsilon(t,\theta)$ | model mismatch, [13] |
| $\kappa$ | condition number, [89] |
| $\mathscr{Z}_N$ | a data set, [12] |
| $\phi_i$ | phase of a signal, [20] |
| $n \times n$ | $n$-inputs $n$-outputs system, [vii] |
| $\alpha_{ji}$ | Fourier coefficients for $j$–input $u_j(k)$, [31] |

| | |
|---|---|
| CF | crest factor, [20] |
| DFT | discrete Fourier transform, [25] |
| DIBS | discrete interval binary sequence, [16] |
| MIMO | multiple-input multiple-output, [vii] |
| MISO | multi-input single-output, [35] |
| MPC | model predictive control, [vii] |
| PF | peak factor, [22] |
| PIPS | performance index for perturbation signal, [23] |
| PRBS | pseudo random binary sequence, [vii] |
| PSD | power spectral density, [18] |
| RBS | random binary sequence, [16] |
| RMS | root mean square, [22] |
| SVD | singular value decomposition, [37] |

The page numbers where the symbols/abbreviations are explained are given in the brackets.

# Index

This thesis, containing 29 figures and 15 tables, was typeset with LaTeX using Bitstream Charter (Charter BT) and Computer Modern type faces. The artwork was prepared using TikZ/PGF (vector & 3D graphics; publishing of final figures to PDF).

A copy of this thesis can be downloaded from the university's website: `www.abo.fi`

Ramkrishna Ghosh

# Input Designs for Identification of Ill-conditined Multivariable Systems

Making an identification procedure efficient enough is an ongoing demand. Typically, designing excitation for ill-conditioned and directional MIMO system is a challenging task. Tailor made input excitation using process knowledge is one way to make the identification more efficient.

The focus of this thesis is to develop a better MIMO identification process with proper design of input excitation signals. A comparative study to choose the proper excitation to identify a higher order ($n \times n$ with $n > 2$) inputs-outputs ill-conditioned system is new. Two different tools are proposed for analyzing the excitation of the outputs.

Ramkrishna Ghosh | Input Designs for Identification of Ill-conditioned Multivariable Systems | 2016