Linear Parameter-Varying Modeling of a High-Purity Distillation Collumn

Ahmad Alrianes Bachnas

Master of Science Thesis

TU Delft Delft University of Technology

Delft Center for Systems and Control

Linear Parameter-Varying Modeling of a High-Purity Distillation Collumn

Master of Science Thesis

For the degree of Master of Science in Systems and Control at Delft University of Technology

Ahmad Alrianes Bachnas

September 14, 2012

Faculty of Mechanical, Maritime and Materials Engineering (3mE) · Delft University of Technology

Abstract

One of the main reasons for the success story of *Proportional-Integral-Derivative* (PID) control in process industries has been the difficulty and complexity of modeling chemical, thermal, and physical phenomena in these systems. *Linear Time-Invariant* (LTI) identification has been found incapable to accurately capture the dynamics in these applications over the entire operating region, while nonlinear identification methods still often result in over-laborious and expensive process modeling tools with a too complex model for control synthesis. The concept of data-driven *Linear Parameter-Varying* (LPV) modeling offers an in-between-solution over LTI and nonlinear identification by describing the signal relations in a linear manner which vary with the operating point of the system. Furthermore, the LPV modeling and control framework is considered to have the potential to become the long-waited solution for the modeling and control problems of the process field. Through the case study of Propane-Propene splitter, which is a commonly used high-purity distillation column, this thesis explores the potential of LPV modeling and identification for process systems. Importantly, the PP-splitter represents a particularly challenging process system due to its *Multiple-Input Multiple-Output* (MIMO) nature and significant nonlinear behavior in the high-purity operating region. The results presented in this thesis shows that, although such a system could impose problems for local LPV identification approaches (identification w.r.t. constant operating conditions), the application of global LPV identification (identification w.r.t. varying operating conditions) has found to be able to identify the dynamics of the system even under severe noise conditions. Moreover, the global methodology presented in this thesis, which is called the LPV *Least-Square Support Vector Machine* (LS-SVM), belongs to an emerging trend of novel approaches which adopt machine learning theories in the system identification. This method is the current state-of-the-art among such learning approaches in the LPV identification framework.

Table of Contents

List of Figures

List of Tables

Acknowledgements

First of all, i would like to express my greatest gratitude to Allah SWT, for the strengths and His blessing in completing this thesis. Afterwards, i would like to thank the following people for their time and support who in one way or another contributed and extended their valuable assistance in helping me to bring this thesis together.

item

- My supervisor, Paul van den Hof, who has provided support and guidance during my thesis project.
- Roland Tóth, my daily supervisor. Throughout my thesis-writing period, he provided encouragement, sound advice, good teaching, lots of good ideas, and importantly the fun on doing this thesis work. I would have been lost without him.
- Ali Mesbah my second daily supervisor, for his knowledge in process he helped to make this thesis possible.
- Jobert Ludlage for helped me to validate the high purity distillation column model.
- My wonderful wife, Anti, who have given up far too much to allow me the time to do this thesis, and for taking the time to proof read this thesis, but mainly for listening to my rants and putting up with me.
- Daya and Bee for their great support and for printing this thesis.
- Bhimo, Oed, Anin, Shintoq, Fany and everyone in PPI Delft who has being great company and great supporter throughout my study.
- Lastly, and most importantly, my parents for their patience and being a great support system for me. They have taught me lots of things and loving me all this time. To them I dedicate this thesis.

Delft, University of Technology
September 14, 2012

Ahmad Alrianes Bachnas

Chapter 1

Introduction

Since the dawn of the petrochemical industry, the products at this economical segment have had a significant role in our daily lives. Imagine life without gasoline, fertilizers, synthetic fabrics, and plastics. All of these products and many more are the products of the petrochemical industry. The massive usage of the petrochemical products is driving the growth of this industry. To sustain this growth, engineers and scientists are continuously working to find a way to develop better products or to increase the economical efficiency of the production. In petrochemical plants, theories from the field of system and control engineering have been applied since the very beginning of this industry. The application of this field have purposes to reach and maintain some important operational variables such as heat, pressure, or volume, which are vital to obtain the desired product quality as well as maintain safe operation of the plant. In order to reach these purposes, system and control engineers in industry have are commonly utilizing *Linear Time-Invariant* (LTI) control-designs. The LTI approaches are preferred as they guarantee high performance and reliability, have easy and quick design schemes, and engineers have a vast experience in their application. In the control design process, to optimally synthesis control schemes for the designated plant, one requires a good model of the plant, in order to accurately describe the behavior of the plant. Which further emphasize the importance of modeling and system identification theories and application.

The demand for accurate and efficient control of today's industrial applications is driving the system identification field to cope with the constant challenge of providing better models of physical phenomena. As this demand increases, the popular LTI system identification approaches reach their limitation. Incase of a strong nonlinearity behavior of the system, the identified LTI model will only be good in a very narrow neighborhood of the used operating point. This limitation forces people to go back to first-principle modeling approaches, which are based on prior knowledge of first principle laws of the designated system. The application of first principle modeling will often results in a very complex models of the system dynamics, and is often found to be very laborious and expensive in practice. Based on the circumstances of LTI identification and first principle modeling, the concept of data-driven Linear Parameter-Varying (LPV) modeling and identification framework, can be a good solution. This is because LPV identification approaches offers as an in between solution over LTI identification and the first principle modeling.

The concept of LPV system originates from an older method known as gain-scheduling. The basic concept of gain-scheduling is to linearize a given nonlinear model from a system at different operating points. Then an individual controller for each local LTI model is designed and interpolated to create a global control solution for the entire operating trajectory. To describe the changes of the operating point, a signal which called the scheduling signal is introduced as a variable on this system. Despite the early popularity of this methods, gain scheduling are still remains as an ad-hoc methodology. Some of its underlying properties such as the robustness, performance, or even nominal stability of a global gain scheduled controller are not addressed explicitly in the design process (Shamma and Athans 1992). In the mid 1990s, LPV control started to get attention when the first results about the extension of $\mathcal{H}_\infty$ and $\mathcal{H}_2$ optimal control through *Linear Matrix Inequalities* (LMIs) based optimization appeared (Packard 1994; Apkarian and Gahinet 1995; Scherer 1996). This was followed by the robust LTI control originated μ-synthesis approaches (Zhou and Doyle 1998). These approaches guarantee stability, optimal performance, and robustness over the entire operating regime of LPV models. Since then, the LPV field has evolved rapidly in the last 20 years and became a promising framework for modern industrial control (Tóth 2008).

The petrochemical-chemical industry will surely gain benefit from the LPV identification framework. This is because of most of the applied control strategies in this field are *Proportional Integral Derivative* (PID) controllers. As a standard control methodology, PID have been applied in this industry since the very beginning. The main reasons for the success story of PID control in process applications has always been the difficulty and complexity of modeling chemical, thermal, and physical phenomena in these systems. The LTI identification is not capable to accurately capture the dynamics in these applications over their entire operating region, while nonlinear modeling methods are still often result in over-laborious and expensive process modeling tools with a too complex model for control synthesis. Lets take a look at the example of the application of *Model Predictive Control* (MPC) in the process industry. The MPC and its nonlinear extensions are deemed to be the successor to the PID controllers due to its potential to optimally control *Multi-Input-Multi-Output* (MIMO) system under various constraints. The performance of MPC depends on the quality of the derived model used for its control synthesis. Hence, high cost to obtain an accurate model for the system, have suppressed the expansion of MPC in this industry. If a low complexity and low cost modeling / identification techniques which can be encouraged in the process industries, the leash that hold the breakout of MPC might be broken. Then, who knows what will happen next in the field of process industry.

To explore the potential of the LPV modeling and identification framework for the petrochemical-chemical industry, this thesis choose a case study of a high-purity distillation column, specifically a Propane-Propene (PP) splitter plant. Distillation column itself, is a common industrial process which separates two or more chemical substances in a chemical mixtures, based on the differences on their boiling points. The quality of the separated substances are the main concern of the column operational purposes. In order to reach high product purity, the column must be built tall with hundreds of trays. If for each trays there are around 3-4 *Differential Algebraic Equations* (DAE's), we can imagine the complexity and the modeling problem which this column poses. Instead conducting a high cost modeling procedure to model such a complexity, the identification approach offered in this thesis rely on the concept of data-driven identification. This approach only requires excitations of the system to identify the system dynamics. Therefore the hardship of modeling chemical, thermal, and physical phenomena

can be avoided. Additional notes on the high purity distillation column, is that this system represents a particularly challenging system due to its nonlinear behavior, especially in the high-purity operating region (Skogestad and Morari 1986,1987,1988; Finco et al. 1989), the LTI based system identification is definitely not a possible solution.

In a recent study, a traditional LPV identification approach is shown to have failed to capture the dynamics of the high purity distillation column (Huang et al. 2012). Similar identification approach (which are based on the local information on the system dynamics) is also given in this thesis. The results of the experiment have shown that such a local perspective approach still have potential to capture the dynamics of the system. However, this approach have a limitation which is based in the quality and the accuracy of the local models that are used in this approach. On the other hand, this thesis offer different identification perspective, where the identification on the system are done directly from the global information of the system. The method used in this global approach is called the LPV *Least Square-Support Vector Machine* (LS-SVM), a current state-of-the-art among the approaches of LPV identification framework, which adopt a machine learning theories in the system identification applications. The original LPV LS-SVM have been modified to be able to identify the high purity distillation column which is both MIMO and a large scale system. In the end, the identification results shown that the LPV LS-SVM method is able to capture the dynamics of a high purity distillation column, even in a severe noise conditions.

1-1 Problem Statement

The big picture of this thesis project is to find an efficient modeling solution for high-purity distillation columns, focusing on a specific Propane-Propene splitter as a case study. Such systems are well known to be highly nonlinear in operating points where the distillates have high purity level. This was described by MIMO system with a large set of differential algebraic equations corresponding a so called large scale system. Due to these properties, they are representative examples of common behaviour of process systems which good benchmark to test the applicability of LPV modeling and identification.

In order to study LPV modeling and identification of this system, a high-fidelity nonlinear model of the PP splitter have been developed. The startup phase is not considered in this model. The reason is that the main purpose of the control design in a distillation column system is to reach desired product quality in the running column rather than to improve the startup phase.

The purpose of the identification study is to develop an efficient (low order but high accuracy) LPV model of the system which is important for the control synthesis purposes. This thesis will try to solve the identification problem of the column from two different perspective of LPV identification approaches which is local approach and the global approach. We are interested in comparing the methodologies and give a fair picture about their advantages and disadvantages. Properties of the used approaches in this modeling problem, experiment design and validation with respect to the applied methods are crucial issues which are going to be explored in this thesis.

1-2 Overview of Content

This thesis is divided into six chapters that starts with this introduction chapter. The next two chapter will serve as a general preliminary chapters before going into the specific methodologies that are presented in this thesis. In the second chapter, a brief overview of the operation and modeling of distillation columns are given. Within this chapter the step by step procedure that we take to build the high-fidelity model of the column is also presented. This includes assumptions as well as the first principle laws that are taken for the modeling purpose. Since the model of the column will be further used as the data generating system for the LPV identification study purpose, we will state the interesting operating regimes of the system, and give overviews of the input and output characteristic around that region.

The third chapter will explain about the overview of the LPV modeling and identification framework. Some of available model structures, as well as basic properties of the LPV model are given here. This is done to give the reader a basic understanding of the LPV modeling and identification framework and to relate the modeling and identification aspect of LPV from the distillation column perspective. This chapter will be closed by presenting available identification approach that are currently studied/available in the literatures. The specific approaches that will be used for the identification of the distillation column will be further elaborated in the next two chapters. The fourth chapters presents LPV identification from the local perspective, while the fifth is from the global perspective.

From the local LPV modeling perspective, the LPV model are derived by interpolating several numbers of identified LTI model which describe the local dynamics of the system on the points of identification. The local LTI models of the system are identified with the help of *Prediction Error Minimizations* (PEM's) based identification. Different interpolation schemes such as output interpolation, input interpolation, and coefficient interpolation schemes along with different interpolation method such as RBF , polynomial, and bilinear based interpolation are presented in this chapter. For each of this schemes we can independently select different interpolation methods. The experimental result will show the effect of different combinations of interpolation schemes and methods with respect to identification of the column. Interesting issues found later in the experiment show that the PEM identification for the distillation column, can be a hard task. If we further use such a linear model for the interpolation, several issues such as stability and oscillations on the LPV model response might occurs.

In the fifth chapter, the basic of LPV modeling in global perspective is given here. The global identification approach that are applied here are a new non-parametrical approach which merge a machine learning methodology called the *Support Vector Machine* (SVM) in the LPV identification settings. Therefore we also give a basic description on the SVM theories as well as the adoption in the LPV settings. The results of this method have shown that this method are promising to be applied in such a complex/large scale system such as the column.

Last but not least, we will summarize the findings in the works of this thesis in the last chapter. This chapter will be concluded with some recommendation that can done for future improvements in the scope of this thesis projects.

Chapter 2

High-Purity Distillation Column

In this chapter, the dynamical aspects of distillation columns will be in the focus of the discussion. We will start with a general overview of distillation column. The column chosen for the case study is a binary column, where only two components are separated. Afterwards, the discussion will be continued with the modeling of such a column. The modeling part will start with assumptions of the modeling, and then continued by a mathematical description of the dynamics of the distillation column. From the assumptions and equations, a high-fidelity model is produced. This model will serve as the data-generating system which will be used in the further study of the LPV identification in the next chapters.

2-1 Overview of the Distillation Column

Distillation is a common chemical operation to separate liquid mixtures. The separation is done with the help of the difference in the boiling points of the components. In the process industry, this unit operation is commonly accomplished in a system called distillation column. This column have a main goal to collect the separated chemical product in the desired purity levels. As a case study we select a Propene-Propane (PP) splitter, which is one of the typical industrial applications of distillation columns. The PP splitter can be categorized as a binary distillation column, since it separates two mixture types. The end product itself is useful for gasoline production, or as raw material for further chemical synthesis. The diagram of a PP splitter is presented in Fig. 2-1.

A common binary distillation column consists of trays stacking above a reboiler and ends with the condenser on the top of it. Somewhere in the middle of the column, on one of the trays, a chemical mixture is being fed into the distillation column. The tray where the feeder is located is usually called the *feed stage* and its location depends on the design of the column itself. Different optimization schemes are commonly used to determines the location of the feed tray (Halvorsen and Skogestad 2000). The feed mixture flows continuously through the whole operation of distillation column, from tray to tray until it reach the bottom of the column. In the bottom of the distillation column, there is a reboiler which produces heat that can

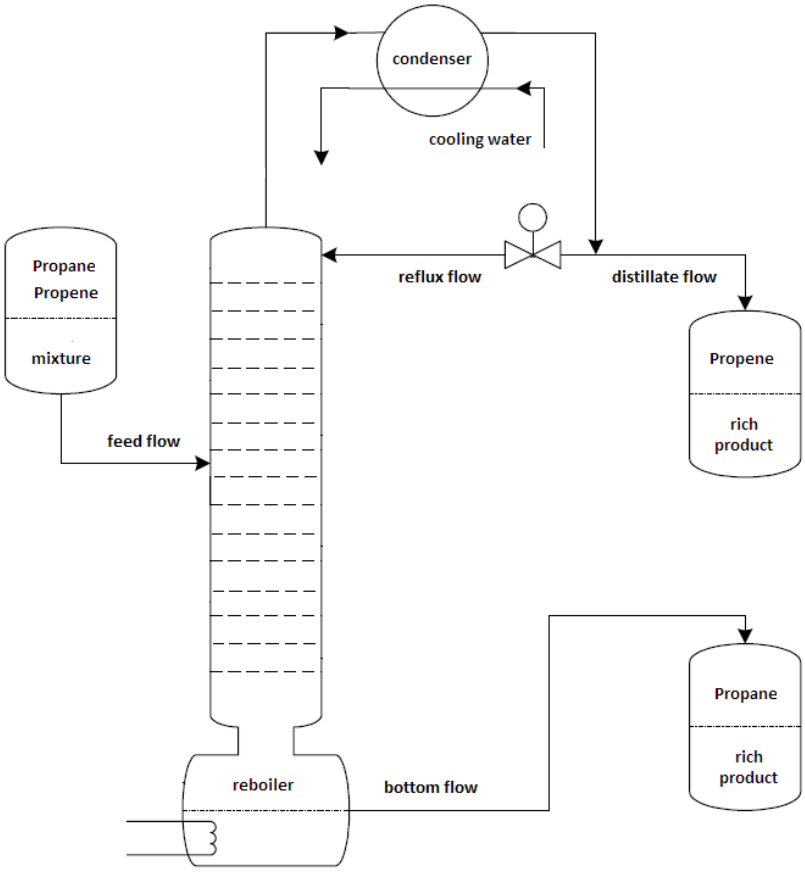


Figure 2-1: The Schematic of a typical PP-splitter (binary distillation column).

vaporize the mixture. Purposely, the heat produced is able to vaporize the light component and leave most of the heavy component as the bottom product. The vapor flow produced by the reboiler still contains both components and go all the way up to the condenser stage by passing through the holes inside each trays passing through the previous liquid from the mixture flows that are still contained in each trays.

The heat sequentially decreases as the vapors travels in the column. In the top trays, the vapor will contain higher percentage of light component due to the difference of the components boiling points. This vapor is then condensed inside a condenser, and the liquid are collected as the top product. Some of the liquid produced by the condensation is pumped back, which is called the *reflux flow*. The trays will hold the liquid mixture with capacity based on the design of the tray. The amount of liquid that can be contained on a tray is called the *molar holdup*. The excess liquid passing the maximum holdup value goes down to the tray bellow, and for the lowest level tray will go to the reboiler tank just as explained before. In these trays, heat and material transfer takes place. The trays bellow the feed stage are usually called the *stripping section*, while above the feed stage are usually called the *enriching section*. This will be explained later in terms of first principles based model.

The parameters of distillation column such as reflux flow, reboiler duty, distillate flow, and/or bottom flow according can be chosen as the input for the control configurations which are

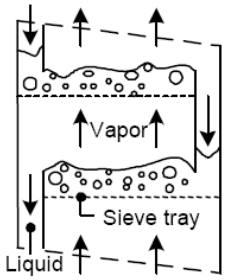


Figure 2-2: Flows inside the trays

used to control the output of the system which is the desired product purity. There are many control configurations that have been implemented and studied on distillation columns. Commonly used control configurations are the control of the liquid-vapor flows, and the control of the bottom-distillate product flows (Skogestad and Morari 1987; Skogestad et al. 1990). A classical controllers such as *Proportional-Integral* (PI) or *Proportional-Integral-Derivative* (PID) controller, can be applied to solve the control problem of the distillate or the bottom product, but have difficulties in controlling both at the same time due to the strong coupling between the input-output channels and the nonlinear behavior of the plant. Another problem in the distillation column occurs when the purity rises to a high level. At this point, the column starts to have a strong nonlinear behavior such as the directionality phenomenon. When the directionality phenomenon occurs, an input of the system may result in different output direction, thus making it hard to find an input combination to produce outputs of high purity.

This nonlinear behavior presents a difficulty for a linear controller due to their limitations. Robust control, is one of the available solutions for this problem (Skogestad et al. 1991), as this control scheme can also ensure stability on multiple operating trajectories. However, by ensuring the stability, it has the drawback of decreased nominal performance. That is why new solutions are need to be explored. The use of the LPV identification framework is motivated for this problem, as it can help in controlling a nonlinear system by providing an accurate model with an attractive computational load. The identified LPV model will keep the linear description of the system and make it easier and more robust to synthesis a controller for the system. To study the feasibility on applying the identification framework for the distillation column, the dynamic behavior of the column need to be studied. For this purpose a high fidelity model have been developed.

2-2 Modeling of a Distillation Column

This section will explain the derivation of a first-principle model for the distillation column setup. The derived model can represent the main dynamical properties of a real-worked plant but at the same time allows simulation of different aspects of the physical behavior of the system. The modeling process starts with stating assumptions related to the model. Afterwards, the first principles equations are used to construct the high fidelity model.

2-2-1 Assumptions

To derive simple analytical expressions of the system behavior in the form of *differential algebraic equations* (DAE's), certain assumptions are needed to be taken. These assumptions are regularly used in practice for the modeling of distillation column. The description of each assumptions are given in the sequel:

- Equilibrium on each trays

  This assumption states that in each tray, the amount of incoming liquid flows (L_{n+1}) and vapor flows (V_{n-1}) are equal to the amount of liquid flows (L_n) and vapor flows (V_n) leaving the trays. Note that there is no chemical reaction taking place on any trays. This means that on each tray, the amounts of vapor and liquid are in equilibrium. One may refine the equilibrium stage concept, by introducing back mixing or a Murphee efficiency factor , but these are expected to have relatively little impact on the considered modeling study (Halvorsen and Skogestad 2000).

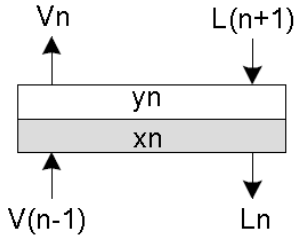


Figure 2-3: Vapor Liquid Equilibrium on tray number n

- Binary Mixture

  This system is assumed to be a *binary mixture*. Although in reality, the feed of a PP splitter usually consists of not just Propene-Propane but also other substances like Ethane, Butylene, Isobutane, and other heavier components. These heavy components usually have relatively low amount in the feed composition (less than 1%), hence they are often negligible. These other substances are usually heavier than both propylene and propane and will propagate to the bottom flow along with propane without any means of control. Based on these, the assumption of a *binary mixture* will hold well on the studied phenomenon.

- Constant relative volatility

  With the assumptions of a *binary mixture* and *equilibrium state*, it is logical to assume a constant relative volatility. The relative volatility (α) is a constant which is assumed to be independent of composition and pressure. The assumption states that on each given liquid molar fraction, the molar fraction of the vapor phase can be obtained with the equation:

$$y_n = \frac{\alpha x_n}{1 + (\alpha - 1)x_n}; \quad \text{for all} \quad i \in 1, \ldots, n \tag{2-1}$$

  where x_n is the mole fraction of the most volatile component in the liquid phase on tray number n, and y_n is the mole fraction of the most volatile component in the vapor phase on tray number n.

Also note that in the case study, propene and propane have a similar characteristics with similar chemical structure, which is also a motivation to use this assumption.

- Constant molar flows

  This assumption indicates that the molar flows between trays do not change from one stage to another (with the exception of the feed tray). This is because there is no feed or a product removal between the stages and the column operates in an equilibrium state:

$$L_n = L_{n+1} = L_{n-1}; \quad V_n = V_{n+1} = V_{n-1}, \tag{2-2}$$

  where L_n is the liquid flow on tray number n and V_n is the vapor flow on tray number n. This assumption holds well since the change in the internal flows (ΔV, ΔL) do not drastically affect the distillate and the bottom product. While the external flows (D/F , B/F) have bigger impact on both the distillate and the bottom product purity (Halvorsen and Skogestad 2000).

- The distillate and bottom flows are ideally controlled

  This means that the change in the liquid and the vapor flows (L and V) have instant effect on both the distillate (D) and the bottom flows (B). Furthermore this would also mean that the liquid and vapor flows can be changed without a delay.

2-2-2 First-principles based Model

In the present study, a basic distillation column model is selected. This model is based on conservation laws and relationships of the variables constituting the equilibrium equation. The energy balance is neglected with assumptions of constant molar flows. Total mass balance is also neglected, since the assumed constant molar flows imply that the liquid holdup remains constant.

Component mass balance:

$$\begin{aligned}
VE_{Top} &: \frac{d(M_d x_d)}{dt} = V_{d-1}y_{d-1} - L_d x_d - D x_d, \\
VE_m &: \frac{d(M_m x_m)}{dt} = V_{m-1}y_{m-1} + L_{m+1}x_{m+1} - V_m y_m - L_m x_m, \\
VE_{Feed} &: \frac{d(M_f x_f)}{dt} = F z_f + V_{f-1}y_{f-1} - L_{f+1}x_{f+1} - V_f y_f - L_f x_f, \\
VE_n &: \frac{d(M_n x_n)}{dt} = V_{n-1}y_{n-1} - L_{n+1}x_{n+1} - V_n y_n - L_n x_n, \\
VE_{Reboiler} &: \frac{d(M_b x_b)}{dt} = L_2 x_2 - V_b y_b - B x_b.
\end{aligned} \tag{2-3}$$

All F, B, D, L, and V have the unit of mol/min. V stands for molar flow of the vapor phase, L is the molar flow of the liquid phase, F is the feed flow, z_f is the mole fraction of the feed, B is the bottom product flow, and D is the top product (distillate) flow. By the equation, the *distillation column* itself can be separated into five parts which are the top tray, top section trays, feed tray, bottom section trays, and the bottom trays. Hence, the trays numbering starts from the bottom and ends at the top $(b, 2, \ldots, n, \ldots, f-1, f, f+1, \ldots, m, \ldots, d)$. Parameter M indicates the molar holdups of the trays, the reboiler, and the condenser distinguished by the indices (the constant molar holdup value is equal to the liquid amount that can be contained in each trays).

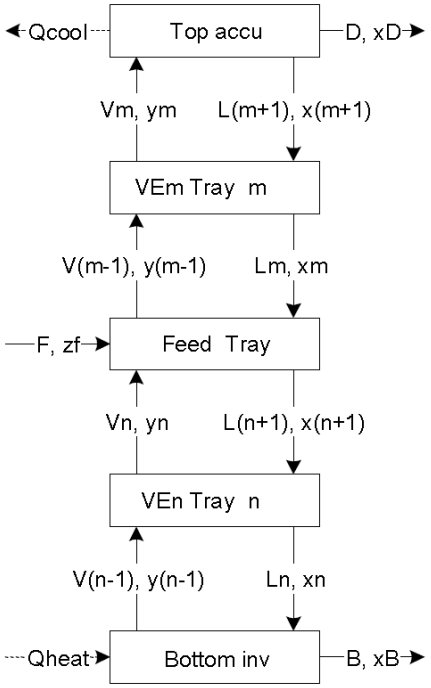


Figure 2-4: Flow diagram of the trays and the accumulators.

Due to that, the vapor and liquid flows are constant on each section of the distillation column:

$$\begin{aligned} V_m &= V_n + (1 - q_f)F, \\ V_n &= V_b, \\ L_m &= L_d, \\ L_n &= L_m + q_f F, \end{aligned} \tag{2-4}$$

where q_f is the fraction of liquid in the feed. In reality, the vapor flow depends on the reboiler heat, and the liquid flow depends on the reflux rate in the distillation tank. However, since we assume that the distillate and bottom flows are ideally controlled, instant change on heat and reflux rate have an instantaneous impact on the vapor and liquid flows. Due to this, the vapor and liquid flows will be used as the input channels for the model, while the output channels will be the top and the bottom products composition, and feed properties such as flow, liquid fraction, and mole fraction will act as disturbances.

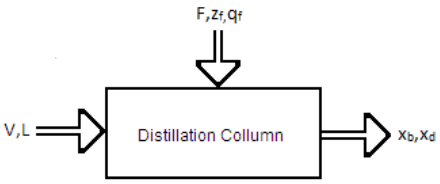


Figure 2-5: Input-Output diagram of the system.

2-2-3 Degree of Freedom Analysis

The model consists of several DAE's. The variables of the system model are:

- N liquid mole fractions (x)
- N-1 vapor mole fractions (y)
- 3 feed properties (F, z_f, q_f)
- 1 liquid holdup (M)

The following equations:

- N component mass balances
- N-1 VLE equations
- 1 distillate flow equation
- 1 bottom flow equation
- 1 total mass balance
- 1 total component mass balance

The degree of freedom is given by:

$$DoF = (2N + 3) - (2N + 3) = 0, \tag{2-5}$$

with the given assumptions and conditions, the degree of freedom of the model is equal to zero, which means by specifying the input signal trajectories V and F„ the trajectory of the output signals, x_d and x_b are uniquely specified.

Table 2-1: PP-splitter parameter

z_F	α	N	N_F	M(kmol)
0.65	1.12	110	39	30

2-3 Nonlinear Model of the Distillation Column

Based on the modeling assumptions and the first principle equation explained above, a high fidelity model have been developed. The model will act as the distillation column setup that we aim to be identified by using LPV data-driven approaches. The parameters of the model are based on the example of a PP-splitter model which was studied in Skogestad and Morari (1988). These parameters are given in Table 2-1:

The column has 110 trays with the feed on the 39th tray (whether this feed location is optimal or not is not the main concern in this thesis). The amount of the liquid holdup in each trays is 30 kmol, while the feed composition is 0.65 with respect to the light component (Propene). The high fidelity model is developed with Matlab Simulink using level-1 S functions to make user defined blocks for the differential equation of the distillation column. The other square shaped blocks refer to different algebraic equations and are essential components to simulate the dynamic behavior of the distillation column.

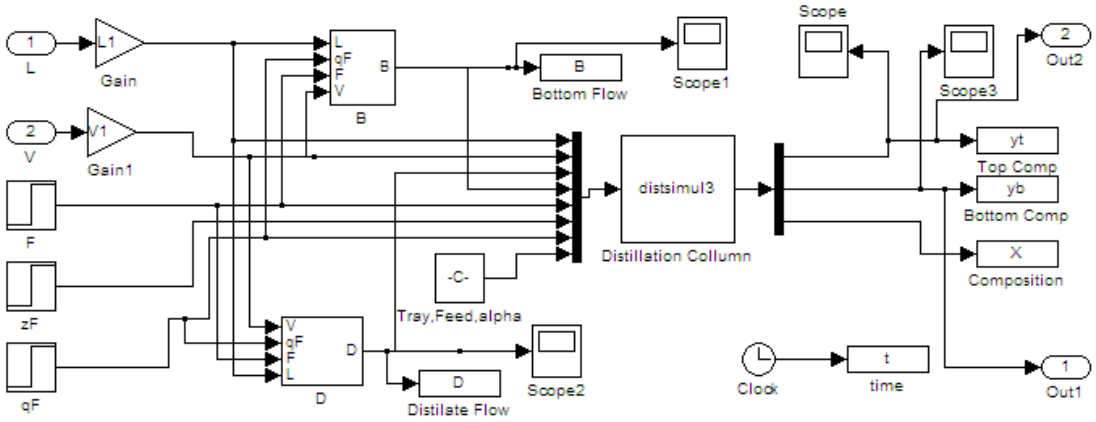


Figure 2-6: The developed non-linear Simulink model.

The Simulink model diagram is described In Fig. 2-6. In this Simulink model, the B and D blocks have algebraic equations which describe the instantaneous change on V, L, F, and q_F, while the ordinary differential equations (Equation 2-3) are located in the *distsimul3* block. This high fidelity model will provide us with a realistic behavior of the studied distillation column, which we can use for studying the applicability of LPV modeling and identification approaches for this system.

2-3-1 Distillation Column Operating Regime

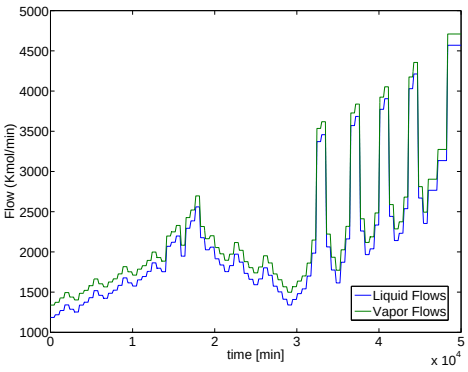
In this high fidelity model there are two input channel which are the liquid flows as the first input and vapor flows as the second input, and two output channel which are the bottom product composition as the first output and top product composition as the second output. As the goal of the distillation column system is to separate a mixture into products with exact purity, we can say that these two output channel characterize the operation of the system. In other word the operating points of the distillation column are denoted by the output of the column. Most of the system, including distillation column, can be operated in several different operating points. The known operational range of the system (which cover multiple operating points) will be further called as the operating regime of the system.

The next tasks that in this thesis, are to apply LPV identification approach to capture the dynamics of the distillation column, especially within its operating regime. Therefore, before we move further into the details of the LPV identification approach, we have to priory decide the interested operating regime of the system. The decision of the system operating regimes have to be made such that it provide a realistic operational aspect of the system. The dynamics of the distillation column, in the high purity region (≈ 0.99 on a product composition) is definitely an interesting operating point that will be covered. However, if the identification can capture not only the high purity region dynamics but also the dynamics on lower purity region, that can increase the usefulness of the identified model. After some operational consideration, we decide to choose 0.85 on both product compositions as the lowest interested composition in the system operating regime. Hence, the operating regime of the distillation column is chosen as [0.85 0.99] for the top product composition, and [0.01 0.15] (since the modeling are done with respect to the composition of Propene as the light component) for the bottom product composition. This range of distillation column operating regime will provide multiple interesting combination of product specification, as well as a challenge in the identification study point of view due to its nonlinear behavior within the chosen range.

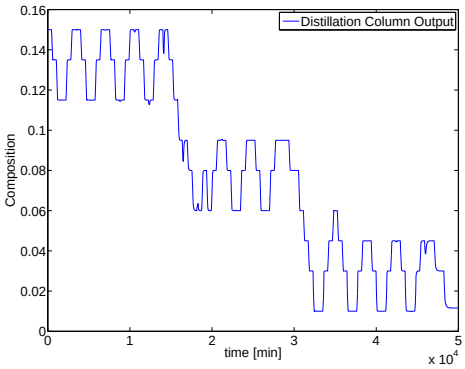
Based on the LPV identification approach that are used in the next chapters, the input channels of the system are used to excite the system to obtain the necessary input-output data sets for either identification, parameters optimization, or validation purposes. An example of the dynamics of the system in within its operating regimes can be seen in Figure 2-7.

This input-output data set will be used further for interpolation weight parameters optimization in local LPV identification approach (See 4-3), and for the identification data set (with added measurement noise) in the global LPV identification approach (See 5-4-1).

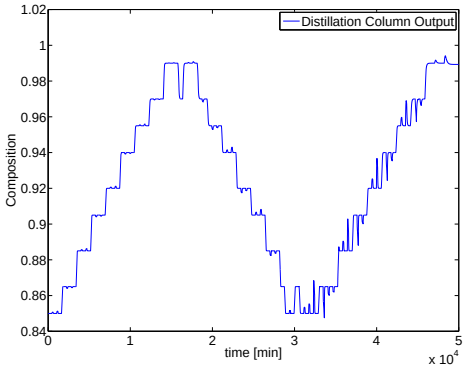
As an additional notes, we will only consider measurement noise of additive white noise for this system, and we are not planning to induces additional disturbances on feed flow and feed component composition. This is because we would want to start with the simple noise description on the high purity distillation column setup. Moreover, on the literature survey stage we can see that there are no resonance or anti resonance peak found in the frequency response of the linearized model (of the high fidelity model) over the operating regime of the system. Therefore, low disturbance amplification is expected in this system.



(a) Distillation column input



(b) Bottom product composition



(c) Top product composition

Figure 2-7: Input-Output data set used for identification.

2-4 Conclusion

High purity distillation column represent a large scale system with high complexity. Due to the unavailability of real distillation column setup and data for the identification study, a high fidelity model of the column have been build to realistically represent the behavior of the system. In a real life case however, the same assumption and modeling procedure used to model this high-fidelity model, might result in less accurate model of a real operating distillation column. The effort and the cost to produce an accurate nonlinear model via first principle modeling can be unimaginably high. Hence, data driven identification approach can be a solution to obtain an accurate model of such a system. However, due to the nonlinearity characteristic of the high purity distillation column, the popular LTI system identification methodology will definitely fail to accurately capture the dynamics of the system within its operating regime. LPV identification in other hand are believed to be able to accurately describe this dynamic while still maintaining low complexity of the model. Before we apply specific identification approach that in the distillation column, we will elaborate more about the basic of LPV modeling and identification framework in the next chapter.

Chapter 3

The Linear Parameter-Varying Framework

The general picture of *Linear Parameter-Varying* system (LPV) is presented in this chapter. As we aim to model the high purity distillation column with LPV identification, it it is necessary to understand the basic of LPV system. First, we will start with the description of LPV system, the scheduling variable and its relation with the input and output signal. The scheduling variable is the main component of the LPV systems which let us keep linear descriptions of a system, but still let the LPV model of the designated system to cover nonlinear dynamic of the system. The available LPV model structures, such as *input-output* and *state space*, will be presented in the sequel. These LPV Model structures are based on LTI model structures, with several modification due the inclusion of scheduling variable. Moreover, the scheduling variable plays an important role on the possibilities as well as the restrictions on LPV model structures transformation. To further support the basic understanding of LPV system, the stability criterions of LPV system are also presented further in this chapter. This chapter will be closed by presenting the available LPV identification approach and also hinting the specific identification approach that will be used for the identification study on the high purity distillation column.

3-1 The LPV System Description

The idea behind the LPV modeling concept is to parameterize a system as a dynamic mapping dependent on the a scheduling signal p. Hence, the relation between the input signals u and the output signals y of the system(i.e., the parameters of the underlying system model) are dependent on p.

The resulting system can be described as a convolution of the scheduling variable p ($p : \mathbb{R} \rightarrow \mathbb{P}$ with $\mathbb{P} \subseteq \mathbb{R}^n$) and the input u. In discrete time, this convolution is given as

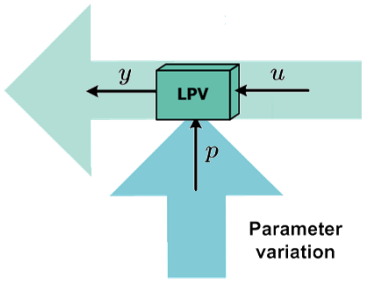


Figure 3-1: LPV system.

$$y(k) = \sum_{i=0}^{\infty} g_i(p,k)u(k-i), \tag{3-1}$$

where each g_i ,or commonly referred as the markov parameters, is the function of the scheduling signal p and the discrete time $k \in \mathbb{Z}$, $u : \mathbb{Z} \to \mathbb{R}^{n_{\mathbb{U}}}$ is the input, and $y : \mathbb{Z} \to \mathbb{R}^{n_{\mathbb{Y}}}$ is the output of the system. The shifted input signal is denoted by $u(k-1)$ which is also commonly written as multiplication of u by the shift operator q, i.e., $q^{-i}u(k) \triangleq u(k-1)$. In Equation 3-1, $\{g_i\}_{i=0}^{\infty}$ defines the dynamical relation between u and y, which in many cases are assumed to have *static dependence* on p. This means that $\{g_i\}_{i=0}^{\infty}$ depends only on the instantaneous value of the scheduling signal. In the case when the scheduling variable is constant, the LPV system description is equal to an *linear time invariant* (LTI) system, as each g_i becomes a constant.

The description in Eq 3-1 is a noise free representation of the output evolution. In practice the real system is a subject of noise and disturbances. If the noise is considered to be an additive noise, the output equation of the whole model can be written as:

$$\tilde{y} = y + v, \tag{3-2}$$

where $\tilde{y}$ is the measured model output, y is the noise free model output, and v is the additive noise which is considered to be a stochastic process. The noise is usually chosen to be white noise or filtered white noise, as in that case it can be modeled and treated in a *prediction error* (PE) fashion. Such an assumption has been found to be applicable for most real-world systems, such as distillation column.

3-1-1 Scheduling variable for high-purity distillation column

The signals chosen for the scheduling variables come directly from the original system. This scheduling variable should be able to display the changing dynamics of the system or known as the changing operating point of the system. As said in 2-3-1 the operating points of the distillation column are usually denoted by the composition of the output (Top or bottom product). This composition can be measured via composition sensor that are placed in both product flows. Other parameters which vary as the operating points of the system changes, such as the composition or temperature measured on other trays can also be used as the

scheduling signal. However, the choice of the scheduling variable should also consider applicability on the column itself. After we have chose a specific variables for scheduling, we have to consistently use it for the whole identification approach, as the choice of the scheduling variables define the possible dynamical behaviors of the produced LPV model structure, hence, the representation capability of the model.

In the next two chapters, the identification of the distillation column is done with two different approaches, which are from the local perspective and from the global perspective. To better show the flexibility of the selection of the scheduling variables, we will use a different scheduling signal in each of those approaches.

3-2 LPV model structure representations

The similarity between LPV and LTI system is one reason to motivate the application of LPV system. As techniques for LTI system are already well known and already implemented in various application. Like LTI systems, an LPV system can be described in terms of different representation forms. System representations such as input-output and state-space forms are examples of LPV system representations. Although, in the LPV case, the transformation between these representation forms is more complicated than in the LTI case (See 3-2-3).

3-2-1 LPV input-output representations

Input-output representations are well-known forms used to describe an LTI system. This structure represent the input-output relation of a system by a set of polynomial equations. Such equations originate from a difference equation based representation of the system behavior. The idea to use an easily identifiable representation ,such as polynomial equations in the difference operator, has been the driving force in the development of LPV-IO identification approaches. The LPV-IO representation in discrete time can be described in the form:

$$y = -\sum_{i=1}^{n_a} a_i(p)q^{-i}y + \sum_{j=0}^{n_b} b_j(p)q^{-j}u, \tag{3-3}$$

where $\{a_i\}_{i=1}^{n_a}$ are coefficients of the output side and $\{b_j\}_{j=0}^{n_b}$ are coefficients of the input side of the dynamical relation, with $n_a \geq 0,\ n_b \geq 0$ and $n_a > 0$. If $n_a = 0$ then there are no output dynamics involved and Eq. 3-3 corresponds to a *Finite Impulse Response* (FIR). These coefficients are often considered to be functions of the instantaneous value of p, i.e., $p(k)$, which is called static dependence. The identification task of LPV-IO representations is to find the coefficients a_i and b_j such that the LPV form written in Eq. 3-3 can represent the dynamical behavior of the considered system.

3-2-2 LPV state-space representations

Similar to the LTI case, LPV systems also have state space (SS) representations. In the LPV case, the state variables are also affected by the parameter variation. The LPV-SS representation in discrete time is commonly written as:

$$\begin{aligned} qx &= A(p)x + B(p)u, \\ y &= C(p)x + D(p)u, \end{aligned} \tag{3-4}$$

where $x : \mathbb{Z} \to \mathbb{R}^{n_{\mathbb{X}}}$ is the state variable and $\{A(p) \in \mathbb{R}^{n_{\mathbb{X}} \times n_{\mathbb{X}}}$, $B(p) \in \mathbb{R}^{n_{\mathbb{X}} \times n_{\mathbb{U}}}$, $C(p) \in \mathbb{R}^{n_{\mathbb{Y}} \times n_{\mathbb{X}}}$, and $D(p) \in \mathbb{R}^{n_{\mathbb{Y}} \times n_{\mathbb{U}}}\}$ are the state-space matrix functions. These matrices (($A(p)$, $B(p)$, $C(p)$, $D(p)$)) are matrix functions with static dependence on the scheduling variable p. In this manner, the scheduling signal p have a freedom to govern each relation between input, output, and states of the system directly. In practice, these SS matrices are hard to estimate, and the estimation is mostly based on approximation. This is because it is difficult to correctly estimate state x without priory knowing the dependence of A, B, C,and D on p However, the SS representation are usually required for most control synthesis to formulate control synthesis efficiently in terms of algebraic relation.

3-2-3 LPV model structure transformations

Each representation has its own advantage compared to the other. When compared with the identification in the LPV-SS forms, the identification in LPV-IO structure is relatively easier to be done due to it is less vulnerable to explosions of data-equations (*curse of dimensionality*) . However, many advanced control synthesis approaches require LPV-SS representation based models. Based on this problem, a method to transform LPV-IO model structure to an LPV-SS structure is required. In the LTI case, this transformation is called state-space realization. However in the LPV case, the inclusion of scheduling signal will effect the state-space realization. Consider a second-order SS representation:

$$\begin{aligned} \begin{bmatrix} x_1(k+1) \\ x_2(k+1) \end{bmatrix} &= \begin{bmatrix} 0 & a_2(p(k)) \\ 1 & a_1(p(k)) \end{bmatrix} \begin{bmatrix} x_1(k) \\ x_2(k) \end{bmatrix} + \begin{bmatrix} b_2(p(k)) \\ b_1(p(k)) \end{bmatrix}, \\ y(k) &= x_2(k). \end{aligned} \tag{3-5}$$

With simple manipulations, this system can be written in an equivalent IO form of

$$\begin{aligned} y(k) = {} & a_1(p(k-1))y(k-1) + a_2(p(k-2))y(k-2) \\ & + b_1(p(k-1))u(k-1) + b_2(p(k-2))u(k-2). \end{aligned} \tag{3-6}$$

The equivalent IO form in Eq. 3-6 is not the same as the LPV-IO form mentioned in Eq 3-3. This IO form contains time shifted scheduling signals $p(k-1)$ and $p(k-2)$ which means that the scheduling dependence of this form is not just a *static dependence* but a *dynamic dependence* instead. In order to handle the dynamic dependence, an algebraic framework known as the LPV *behavioral approach* can be used (Tóth et al. 2011). For the formulation of the *behavioral approach*, the LPV-IO and LPV-SS model structures have to be allowed to depend on many time-shifted forms of $p(k)$, i.e., $\{\dots, p(k-1), p(k), p(k+1), \dots\}$. The property change on the scheduling dependency will increase the complexity of the produced LPV model structures, which may requires further modification on the controller synthesis and the hardware implementation of the designed controller.

The forms that we choose for the modeling of distillation column is the LPV IO forms. This was chosen because the complexity of the distillation column might give us trouble on the

explosion of data equation when we are using LPV-SS forms. Moreover, the identification based on a LPV-IO structure is more appealing since it has low computational load and extensive stochastic understanding compared to LPV-SS forms based methods.

3-3 Stability of LPV systems

In this section, the stability concept for LPV system is described. Generally, stability is an important issue when we want to operate, or to model a system. We would always want to operate the column in the safe operational procedure. Moreover, the identified LPV model of the column should give accurate description on the stability of the column. There are several stability concept for LPV systems. In the LPV-IO form the stability criterion includes some concept such as dynamic stability and input-output stability. These two stability concept itself can be divided into two categories. The first category is stability with respect to constant (or slowly changing) scheduling variable p which is called *frozen stability*. And the second category is the stability with respect to varying scheduling variable p called *global stability*. Although the frozen stability does not guarantee global stability, it is still useful to be considered as the first stability test. If a system is proven to be unstable under constant or small changing scheduling trajectories, then the test on varying scheduling variable is not necessary since the system will be also globally unstable.

3-3-1 Dynamic Stability

Dynamic stability is the most general aspect of stability. The concept of dynamic stability is formulated around the autonomous part of the system, where there are no input signals given to the system. A system is called dynamic stable, if the dynamic response of the system under small perturbation remains close to its equilibrium value. For an LPV system, this stability term have to be achieved for any arbitrary scheduling variable trajectory. For the distillation column system, without any perturbation on its inputs, the column will remains close to their current the equilibrium value which is the operating point inside the operating regime of system. The dynamic stability criterion can be written as:

$$\| y(k,p) \| \leq \varepsilon \quad \text{for all} \quad k \geq 0 \quad \text{and} \quad p \in \mathbb{P}^{\mathbb{Z}}, \tag{3-7}$$

with $y(k,p)$ is the system output in an arbitrary norm ($\| \cdot \|$). This definition comes from the properties of linearity (the only fixed point of the dynamic relation is 0) and time-invariance (stability on $k \geq 0$ implies stability on $k \geq k_0$ for all $k \in \mathbb{N}$) of the LPV system class. Another notion is *asymptotic stability*, this definition not only requires the system response to be near of the vicinity of the equilibrium but also requires the system response to return to its equilibrium value when time goes to infinity ($y(k,p) \to 0 \quad \text{as} \quad k \to \infty$).

3-3-2 BIBO Stability

The definition of *Bounded-Input Bounded-Output* (BIBO) stability can be understood with respect to LPV systems as well. An LPV dynamical system $\mathcal{S}$ is defined to be BIBO stable,

if all bounded input trajectories will result in bounded output trajectories. In ℓ_τ norm with $1 \leq \tau < \infty$, the definition of BIBO stability with respect to an LPV system reads as:

$$\sum_{k=0}^{\infty} \parallel u(k,p) \parallel_{\ell_\tau} < \infty \quad \Longrightarrow \quad \sum_{k=0}^{\infty} \parallel y(k,p) \parallel_{\ell_\tau} < \infty, \tag{3-8}$$

while in the ℓ_∞ norm, the condition is:

$$\sup_{k\geq 0} \parallel u(k,p) \parallel_{\ell_\infty} < \infty \quad \Longrightarrow \quad \sup_{k\geq 0} \parallel y(k,p) \parallel_{\ell_\infty} < \infty, \tag{3-9}$$

where $u(k,p)$ is the system input and $y(k,p)$ is the system output. Note that an LPV system is globally BIBO stable if it is BIBO stable for any arbitrary scheduling variable trajectory. Asymptotic stability implies global BIBO stability.

For the distillation column however, the BIBO stability have to be taken cautiously due to the limitation the input excitation that can be used for the system. Namely it is not allowed to change rapidly between operating conditions as the vapor and liquid flows in the column will start to significantly vary and possibly collapse, resulting in a turbulent unwanted behavior.

3-4 LPV identification approaches

In general, the procedure on conducting LPV identification approach can be done in two different ways, which is from local perspective, and from the global perspective. The LPV identification on the local perspective are done by utilizing local information of the system, while in the global perspective, the identification of the LPV model, is based from single data sets which contains the system dynamics in its operating regime. As stated before at 3-2-3, the LPV model structure that we are going to use is the LPV IO forms. Therefore, the identification approach that will be explained in the sequel are methodology that is based on LPV-IO forms.

For additional information, although any LPV-SS based identification methodologies will not be described in this section, there are also many LPV-SS forms based identification that can be found in literatures. Some of the examples of LPV-SS based methodologies includes interpolation approaches (which is similar to the IO case), gradient based methods (Lee and Poolla 1996,1997 ; Lee 1997), set-membership approaches (Sznaeier et al. 2000), and subspace techniques (Verdult 2002;Verdult and Verhaegen 2002,2005; Wingerden and Verhaegen 2008).

Local identification perspective

These methods apply the gain-scheduling concept where local (*frozen*) dynamics of the system are identified first and then interpolated to produce the LPV model. What means by local dynamics are the linear dynamics of the system in points within the operating regime of the system. By using this limited information in the form of linear model output response or their coefficient, the LPV identification approach will try to capture the system dynamics in the system operating regime. The interpolation can be applied in the input side of the local

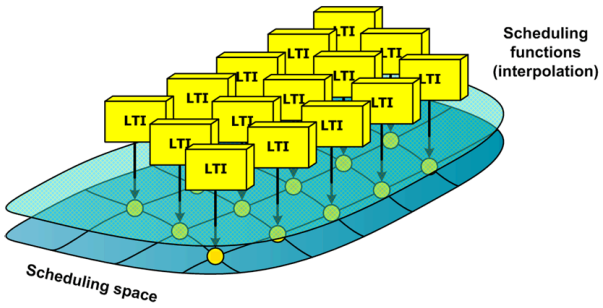


Figure 3-2: Local LPV identification approach.

models (Zhu and Ji 2009), output side of the local models (Zhu and Xu 2008), or directly in the coefficients (Martinez et al. 2009; Bolea et al. 2007).

Recently, a study on this approach for the distillation column system have just been published (Huang et al. 2012). This literature show that the LPV identification on the local approach, can become very challenging. As at the high purity region the nonlinear behavior of the column become very strong, the local identification approaches find difficulties to capture the dynamics of the system.

In this thesis we will conduct the local LPV identification approach with different configuration of interpolation which is at the input side, output side, and at the coefficient of the local models. In each of this scheme we will use different interpolation method to study the effect on the modeling of the distillation column. Moreover, we found that if the identified local dynamics of the system is accurate, this local identification approach are still able to accurately describe the varying dynamics of the system on its operating regimes (See Section 4-3-1). However it is also found that to obtain a local description on the system using PEM identification techniques can become very difficult (See Section 4-3-2), and the modeling error of the local model can have a significant impact on the quality of the resulting LPV model (See Section 4-3-3).

Global identification perspective

Approaches that fall in this category identify the dynamics of the system directly from single input-output dataset of the system within its operating regimes. One of the notorious ways on doing this is to utilize the linearity-in-the-parameters properties of the LPV-IO representation form. This property allows to use linear regression for the estimation of coefficient a_i and b_j under the assumption of additive white noise (Bamieh and Giarré 1999,2000; Wei and Del Re 2006). The corresponding structure is none other than an *auto-regressive* model with *exogenous input* (ARX). Under different noise conditions, instrumental variable methods have been recently introduced for the estimation of LPV-IO models with output-error and Box-Jenkins type of noise models (Laurain et al. 2010). Extensions of the linear regression method even use statistical sparse estimation methods such as the *Non-Negative Garotte* (NNG) and the *Least Absolute Shrinkage and Selection Operator* (LASSO) (Tóth et al. 2009).

Another global identification methods estimate the LPV model of the system by taking non-parametric manner such as *dispersion function approach* (Hsu et al. 2008) and *least square-*

Support Vector Machines (LS-SVM) method (Tóth et al. 2011). In these approaches, the LPV model coefficient value is estimated at each given values of the scheduling variables. The advantages of this approach is that it does not require priori knowledge of the underlying dependencies. The *dispersion function approach* is able to capture the "staticness" of linearly interpolated map of input output data. However, the larger number of parameters to be estimated implies that the achievable bias-variance trade-off by the dispersion approach is inherently worse than in the SVM case (Tóth et al. 2011). Furthermore, the dispersion method corresponds to a quadratic optimization problem which can be considered to be computationally more demanding than the analytical SVM solution.

The global LPV identification approach that will be conducted in this thesis is the LPV LS-SVM method. This choice is motivated by the complexity and the large scale problem of the high purity distillation column. In the real life setup, it would be very hard to guess which nonlinearities play an important role in this column. The original method are modified in certain parts for the distillation column identification purposes (See Section 5-3-2). Furthermore we have found nice results which show the ability for this method to be applied in the distillation column even with severe noise conditions (See Section 5-4-1).

3-5 Conclusion

The basic knowledge of LPV system have the same importance as the information on the distillation column. It will provide the starting ground for more advanced approach that will be given in the latter chapter on this thesis. In the LPV system framework, the modeling of the underlying system still allow us to keep its linear description. This is very advantageous since we can still adopt nice properties from LTI system framework that have been studied and applied for a long time.

The further works of this thesis will be done under LPV-IO model structures. This choice was motivated because the complexity of the distillation column might give us trouble on the explosion of data equation when we are using LPV-SS forms. Moreover, the identification based on a LPV-IO structure is more appealing since it has low computational load and extensive stochastic understanding compared to LPV-SS forms based methods.

Due to the diversity of the approach that will be studied in this thesis, it is important to investigate the benefits and disadvantages of these approach with respect to each other and with respect to identification of high purity distillation column.

Chapter 4

Local LPV Identification of High-Purity Distillation Column

The purpose of local LPV identification approach is to capture the dynamics of the distillation column by interpolating several LTI models identified at different points of the operating region of the system. These models can be identified by conducting a *prediction error minimization* (PEM) based identification procedure. The interpolation of these models depends on the combination of the interpolation schemes and the interpolation method. The interpolation schemes that will be given in this thesis are output interpolation, input interpolation, and coefficient interpolation. While the interpolation methods are *Radial Basis Function* (RBF), polynomial, and bilinear interpolation. The benefits and the shortcomings of this local approach on the modeling of the high-purity distillation columns will be explored further in this chapters.

4-1 LPV identification from local perspective

From a mathematical point of view, interpolation is a method of constructing new data points within the range of a discrete set of known data points. In the context of local LPV identification approach, the interpolation is applied on the identified LTI models only on chosen points within the system operating regime and have a purpose to describe the global dynamics within the system operating regime. The contribution of each identified LTI models depends on the scheduling variables that are incorporated as the combination of interpolation scheme and interpolation method. However, the selection of interpolation schemes and function, as well as the number of LTI models needed to cover the system operating region, still remains to be decided by the user and hence in the application of these approaches, some ad-hoc design steps are always included. The system information, such as the feasible system input as well as the parameters which represents an operating point of the system, can give a hint about the optimal selection of the schemes, methods, and numbers of LTI model.

Things to note is that these LTI models are only capable to describe the local changes and one can expect that fast changes between the operating points will be dominated by the transient switching from one local model to the other. This are dynamical behavior which is generally not represented by local identification based LPV models. Moreover, the quality of the LPV model will be dominated by the quality of the identified LTI model. The problem occurred in the LTI model such as sensitivity and oscillation behavior will be conveyed to the LPV model (See 4-3-3).

4-1-1 Interpolation scheme

Basically, there are three known interpolation schemes for the local LPV identification approach. Two of them directly utilize the identified LTI models by summation of their outputs, while the other interpolates the dynamical representations of these models. Since the distillation column is a *Multi-Input-Multi-Output* (MIMO) system with two output and two input, each LTI model output $\hat{y}_n(k)$ and input $\hat{u}_n(k)$ contains two components ($\hat{y}_n(k) = [\hat{y}_{n,1}(k) \quad \hat{y}_{n,2}(k)]^\top$; $\hat{u}_n(k) = [\hat{u}_{n,1}(k) \quad \hat{u}_{n,2}(k)]^\top$). These LTI models operate on the deviation variables, and only valid in the range of its operating point. Therefore, the inputs and the outputs of the LTI models need to be trimmed with a trimming values which are associated with the steady state value of the LTI models. Both of the input and output trimming value are a matrix with two components, the input trimming value is given by $u_{\text{trim}_n} = [u_{\text{trim}_{n,1}(k)} \quad u_{\text{trim}_{n,2}(k)}]^\top$, while the output trimming value is defined as $y_{\text{trim}_n} = [y_{\text{trim}_{n,1}(k)} \quad y_{\text{trim}_{n,2}(k)}]^\top$.

Output Interpolation

In the output interpolation scheme, the interpolation function α_n are located on the output sides of the LTI models. These functions, or also known as the interpolation weights, are illustrated as the triangle blocks on Fig. 4-1.

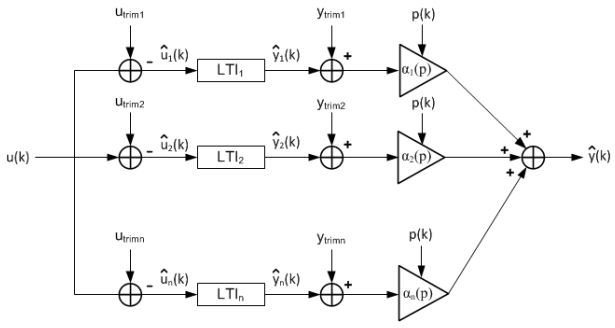


Figure 4-1: Output Interpolation scheme

In this scheme the interpolation weights regulate the proportion of the LTI model output $\hat{y}_n(k)$ that will be added together to produce the LPV system output $\hat{y}$(k). The trimming value on each LTI models are based on their associated operating points. The input trimming values u_{trim_n} subtract the original input sets $u(k) = [u_1(k) \quad u_2(k)]^\top$ which is none other than the liquid flows ($u_1(k)$) and the vapor flows ($u_2(k)$):

$$\hat{u}_n(k) = u(k) - u_{\text{trim}_n}, \tag{4-1}$$

while for the output side, the output trimming value y_{trim_n} add back the steady state value of the LTI model. The LPV model output for this output interpolation scheme is formulated as:

$$\hat{y}(k) = \sum_{i=1}^{N_{\text{LTI}}} \alpha_i(p(k))(\hat{y}_i(k) + y_{\text{trim}_i}), \tag{4-2}$$

with $\hat{y} = [\hat{y}_1(k) \quad \hat{y}_2(k)]^\top$ is the LPV model output, $\hat{y}_n(k) = [\hat{y}_{n,1}(k) \quad \hat{y}_{n,2}(k)]^\top$ the output of LTI model number n, and α_n is the interpolation function. The interpolation weight is not a matrix, hence both output channel $\hat{y}_n(k)$ will be multiplied by the same weighting value.

Input Interpolation

The input interpolation is similar with the output interpolation in the sense of that the output of the derived LPV model directly use the output of each identified LTI model. The difference between this scheme and the output interpolation scheme is that instead of having the interpolation on the output side of the LTI models, the interpolation weights are located on the input side of the LTI models. In this way we expect different behavior due to the fact that we already limit each of the LTI model input by the interpolating weight.

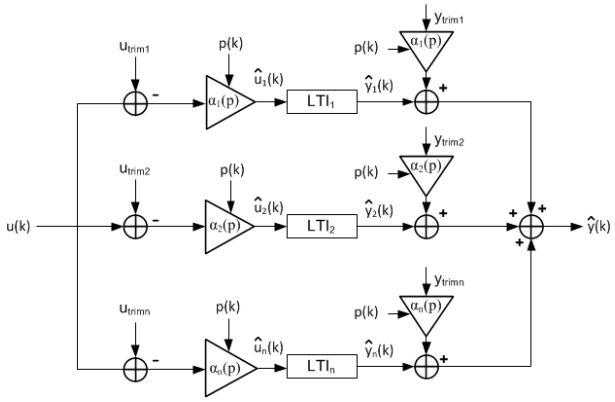


Figure 4-2: Input Interpolation scheme

The interpolation weights $\alpha_n(p(k))$ will decide the amount of the trimmed input

$$\hat{u}_n(k) = \alpha_n(p(k))(u(k) - u_{\text{trim}_n}), \tag{4-3}$$

that is fed into the input of each LTI model. In this scheme, the same interpolation weight $\alpha_n(p(k))$ that is used to weight the trimmed input, will also be placed in the output trimming side. In this way, an LTI model which have least contribution towards the system dynamics will also have least trimming value that will contribute to the LPV model output. The output of the LPV model with this interpolation scheme is defined as:

$$\hat{y}(k) = \sum_{i=1}^{N_{\text{LTI}}} \hat{y}_i(k) + \alpha_i(p(k)) y_{\text{trim}_i}. \tag{4-4}$$

Coefficient interpolation

The coefficient interpolation scheme is a bit different if compared with the previous two interpolation schemes. This schemes utilize only the model coefficient values of the identified LTI models. The interpolation of the LTI models are done directly on their model coefficients, and also on their input and output trimming value. These LTI models parameters vary according to the variation of the scheduling variable $p(k)$. In this manner, the LPV model produced by this interpolation scheme is none other than a linear model which have varying parameters that corresponds to the variation of the scheduling signal $p(k)$.

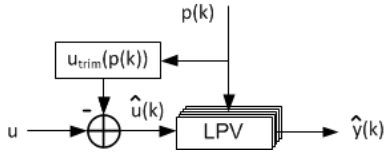


Figure 4-3: Coefficient Interpolation scheme

The way the input-output trimming values used in the LPV system are also different from the previous two schemes. On the input side, the original input $u(k)$ will be trimmed with a varying trimming values

$$\hat{u}(k) = u(k) - u_{\text{trim}}(p(k)). \tag{4-5}$$

For the output side, the output trimming values are introduced directly in the LPV system formulation, via additional trimming dynamics. This additional trimming dynamics are specific for each output channel.

Since we are aiming to identify MIMO system, the LPV system derived via coefficient interpolation scheme should also be formulated in the MIMO forms. To simplify the formulation of the LPV system, we can remove the coupling of each output by transforming the LTI model used for this interpolation into different representations. The new LTI model representations we aim is the one that have same pole on each of the inputs channel. This way each of the LPV system output can be formulated as a series expansions of:

$$\begin{aligned} \hat{y}_n(k) \quad = \quad & -\sum_{i=1}^{n_a} a_i(p(k)) q^{-i} \hat{y}_n(k) + \sum_{m=1}^{n_u} \sum_{j=0}^{n_b} b_{j,m}(p(k)) q^{-j} \hat{u}_m(k) \\ & \underbrace{+ y_{\text{trim}_n}(p(k)) - \sum_{n=1}^{n_a} a_i(p(k)) q^{-i} y_{\text{trim}_n}(p(k))}_{\text{output trimming dynamics}} \end{aligned} \tag{4-6}$$

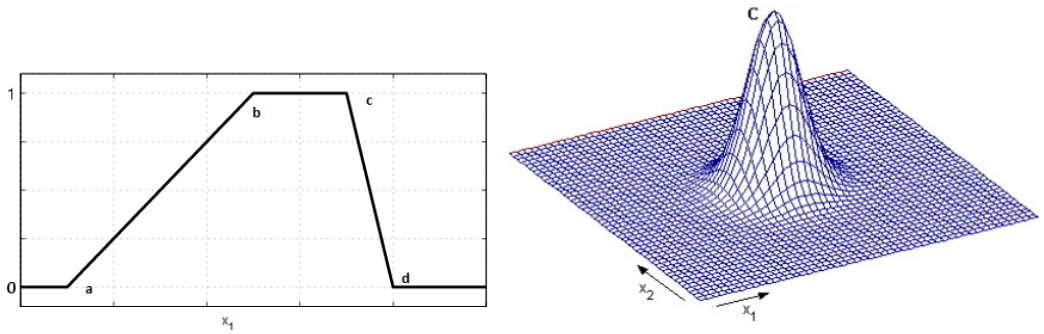
where $a_i(p(k))$ is the model coefficient on the output side and $b_{j,m}(p(k))$ is the model coefficient for the input side, $\hat{u}$ is the trimmed input, and y_{trim_n} is the output trimming value.

4-1-2 Interpolation Method

The interpolation method that are applied on the interpolation schemes can be selected independently. For input and output interpolation, the interpolation method determines the how each interpolation weights are formulated, while for the coefficient interpolation this method determines how to construct the varying model coefficient of the LPV system. This theses will consider three interpolation methods which cover the basic interpolation schemes which can be applied in the local LPV identification approach. They are *Radial Basis Function* (RBF) interpolation, polynomial interpolation, and bilinear interpolation. Each of the interpolation method will have a unique result on the interpolation schemes (See Section 4-3). Additionally, each of this three different interpolation method have different flexibility to assign the parameters of the interpolation function. For polynomial interpolation, the parameters of the polynomial function used for the interpolation can only be selected via solving an optimization problem (Equation 4-10), the bilinear interpolation does not require this step since the interpolation are done in a straightforward manner (Illustrated in Table 4-1), while for the RBF interpolation, this method have the options on to assign the parameters manually, or by solving parameters optimization problem to optimally assign the parameters. By choosing to optimally assign the interpolation weight, we can get better accuracy on the resulting LPV model even with the cost of additional complexity to solve the optimization problem.

Radial Basis Functions based Interpolation

As the name implies, a *radial basis function* (RBF) based interpolation is an interpolation method which utilizes an RBF as interpolation function. The RBF itself is a real-valued function which have a maximum value on its center and a gradually decreasing value as the function variable moves away from its center.In RBF based interpolation, each RBF weights will have a maximum value of one in the center of the RBF and minimum value of zero after some distance from their center. The center of each weights correspond to the operating points of the LTI model which is associated with the interpolation weight. This would mean that on the center of the RBF, only one LTI model (or its parameters in the context of coefficient interpolation) that will contribute towards the LPV system outputs.



(a) 1-D trapezoidal RBF **(b)** 2-D gaussian RBF

Figure 4-4: Example of RBF function.

Some functions such as bell shaped function, gaussian functions, trapezoidal functions, triangular function, etc. can be considered as an RBF. In the case of distillation column, we consider two scheduling variables that act as the variable of the RBF. Intuitively, this would lead to a three dimensional RBF such as a cone shaped function, pyramid shaped function, 2-D gaussian function, etc. However, the optimization to acquire the optimal RBF parameters for more than 1 dimensional parameters corresponds to a difficult optimization problem in general with many local minima and high numerical sensitivity. The optimization can lead to an oddly shaped total weight surface, with some areas on the operating region to be poorly described (have a low total weight) while the other areas might have exceeded the maximum total weights value of one.

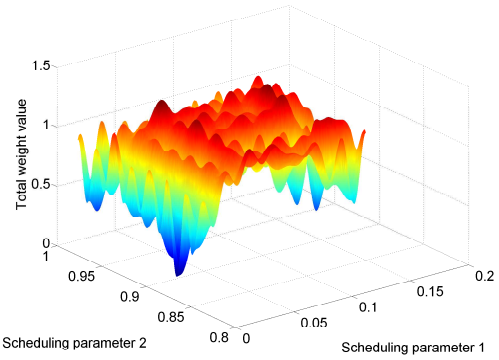


Figure 4-5: Example of a bad total weight surface using 2-D RBF.

The solution to this problem is to separate the scheduling variables into multiple single variable RBF and solve the parameter optimization separately on each scheduling variable. The total weight on a point in that region would be equal to the multiplication of the RBFs which intersect on that point. This is quite similar with *inference* in fuzzy logic which is illustrated in Fig. 4-6.

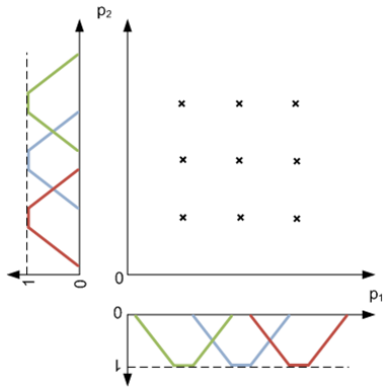


Figure 4-6: Two scheduling variable case.

In this thesis, the investigations are restricted to the trapezoidal function based RBF:

$$\alpha(p; a, b, c, d) = \max\left(\min\left(\frac{p-a}{b-a}, 1, \frac{d-p}{d-c}\right), 0\right), \tag{4-7}$$

where p is the scheduling variable, a, b, c and d are the points of the trapezoid (Figure 4-4.a). The choice of trapezoidal RBF is made through several identification studies in the distillation column setup. Through this studies, we arrive at the conclusion that the trapezoidal RBF have more flexibility to assign the RBF parameter if compared to the other RBF's such as gaussian, gaussian bell, or triangular shape function. This flexibility will give more degree of freedom to optimally assign the parameters with respect to the parameter optimization given in the sequel. Additionally, this became even more beneficial if the location of the identified LTI models are not equally distributed within the system operating regime.

In order to optimally select the parameters of the RBF we can use a nonlinear optimization with additional constraints on the total weight on each of the scheduling variable region:

$$\begin{aligned} \alpha(p_1, p_2; a, b, c, d) \quad = \quad & \arg\min \quad \| Y - \hat{Y} \|_2, \\ & \text{s.t} \quad \sum_{i=1}^{n_{\alpha_1}} \alpha_{1,i}(p_1) = 1 \quad \forall \quad p_1 \in \mathbb{R} \\ & \qquad \sum_{i=1}^{n_{\alpha_2}} \alpha_{2,i}(p_2) = 1 \quad \forall \quad p_2 \in \mathbb{R} \end{aligned} \tag{4-8}$$

where $Y = \begin{bmatrix} y(1) & y(2) & \dots & y(N) \end{bmatrix}^\top$, and $\hat{Y} = \begin{bmatrix} \hat{y}(1) & \hat{y}(2) & \dots & \hat{y}(N) \end{bmatrix}^\top$ are the original system output and the output of the LPV system in the optimization data set $\mathcal{D}_O = \{y(k), u(k), p(k)\}_{k=1}^N$, and $\alpha_{1,n}$ and $\alpha_{2,n}$ are the interpolation functions for LTI model number n with respect to the first and second scheduling variable. This optimization will find the optimal parameters by minimizing the difference between the LPV model output and the original system output, while still keeping each point on the scheduling variable regions to be well weighted. The result of this optimization is very dependent on the optimization data set. Therefore the identification data set have to be designed to achieve the optimal weight parameters (See Figure 2-7).

Polynomial Interpolation

In the polynomial interpolation, basically we fit an n-th order polynomial as the interpolation function. Since we consider two scheduling variables to describe the dynamics of the distillation column, the polynomial function that are used for this interpolation method is a two variables based polynomial function. The polynomial function can be formulated as:

$$\alpha(p) = \sum_{j=1}^{n} \sum_{i=1}^{n} \theta_{(i,j)} p_1(k)^{(n-i)} p_2(k)^{(n-i)} \tag{4-9}$$

with $\theta \in \mathbb{R}^n$ are the parameters of the polynomials, and p_1 and p_2 are the first and the second scheduling variables. Basically, in input/output interpolation, the parameters collected in theta are obtained by minimizing the difference between the output of the original system

$Y = \begin{bmatrix} y(1) & y(2) & \dots & y(N) \end{bmatrix}^\top$, and the output of the LPV model $\hat{Y} = \begin{bmatrix} \hat{y}(1) & \hat{y}(2) & \dots & \hat{y}(N) \end{bmatrix}^\top$ in the optimization data set $\mathcal{D}_O = \{y(k), u(k), p(k)\}_{k=1}^N$:

$$\hat{\theta} = \arg\min \parallel Y - \hat{Y} \parallel_2 . \tag{4-10}$$

This optimization problem can be solved as a least square problem. The reformulation of the problem is done by constructing a regressor:

$$\begin{bmatrix} \varphi_{(1,1)}(k) \\ \vdots \\ \varphi_{(n,1)}(k) \\ \varphi_{(1,2)}(k) \\ \vdots \\ \varphi_{(n,n)}(k) \end{bmatrix}^\top = \begin{bmatrix} p_1(k)^{n-1} p_2(k)^{n-1} \\ \vdots \\ p_2(k)^{n-1} \\ p_1(k)^{n-1} p_2(k)^{n-2} \\ \vdots \\ 1 \end{bmatrix}^\top \tag{4-11}$$

and revoking the optimization as:

$$\begin{bmatrix} y(1) \\ \vdots \\ y(N) \end{bmatrix} = \begin{bmatrix} \varphi_{(1,1)}(1)\hat{y}(1) & \dots & \varphi_{(n,n)}(1)\hat{y}(1) \\ \vdots & \ddots & \vdots \\ \varphi_{(1,1)}(N)\hat{y}(N) & \dots & \varphi_{(n,n)}(2)\hat{y}(N) \end{bmatrix} \begin{bmatrix} \hat{\theta}_{(1,1)} \\ \vdots \\ \hat{\theta}_{(n,n)} \end{bmatrix}. \tag{4-12}$$

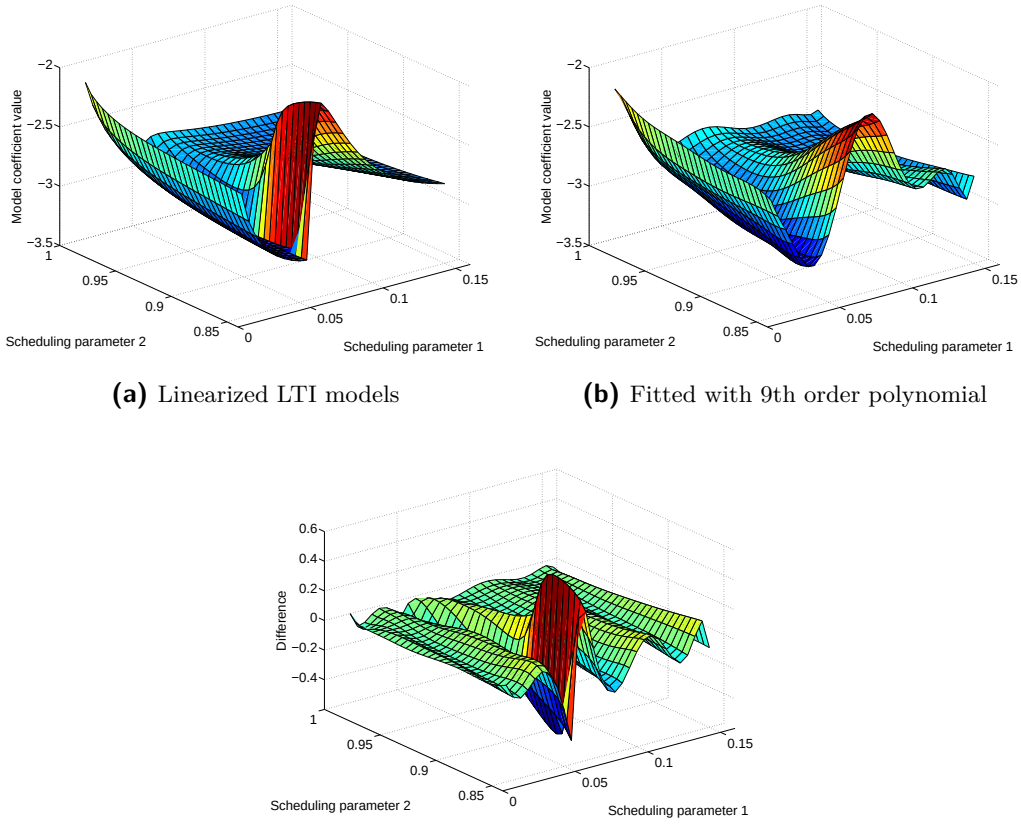
For the coefficient interpolation, this interpolation method uses an n-th order polynomial function (See Eq. 4-9) to fit a polynomial surface through for each one of the model parameters (both input-output trimming values and the model coefficients) on each identified LTI models in the operating regime of the system. An example of this polynomial function fitting can be seen in Fig. 4-7.

Bilinear Interpolation

The last interpolation method which is considered in this thesis is bilinear interpolation. This straightforward method can be described as a *lookup table* method, since it calculates a value (LTI model outputs / parameters) in an operating point based on the other known values at the surrounding points. An illustration of this interpolation method can be seen in Table 4-1:

Table 4-1: Illustration of bilinear interpolation

	scheduling variable 1		
scheduling variable 2	$p_{1,j-1}$	$p_{1,j}$	$p_{1,j+1}$
$p_{2,i-1}$	ϕ_1		ϕ_2
$p_{2,i}$		$\phi_{(p_{1,j},p_{2,i})}$	
$p_{2,i+1}$	ϕ_3		ϕ_4



(a) Linearized LTI models

(b) Fitted with 9th order polynomial

(c) Fitting gap

Figure 4-7: Model coefficient surface example.

To calculate the value $\phi_{(p_{1,j},p_{2,i})}$ which corresponds to the scheduling variables points of $p_{1,j}$ and $p_{2,i}$, we can use a simple arithmetic calculation based on the surrounding values $\phi_1, \ldots, \phi_4$ and their corresponding scheduling variables value. The calculation can be performed as:

$$\begin{aligned}\phi(p_{1,j},p_{2,i}) \quad = \quad & \frac{\phi_1}{|p_{2,i+1}-p_{2,i-1}||p_{1,j+1}-p_{1,j-1}|}|p_{2,i+1}-p_{2,i}||p_{1,j+1}-p_{1,j}| \\ & + \frac{\phi_2}{|p_{2,i+1}-p_{2,i-1}||p_{1,j+1}-p_{1,j-1}|}|p_{2,i+1}-p_{2,i}||p_{1,j}-p_{1,j-1}| \\ & + \frac{\phi_3}{|p_{2,i+1}-p_{2,i-1}||p_{1,j+1}-p_{1,j-1}|}|p_{2,i}-p_{2,i-1}||p_{1,j+1}-p_{1,j}| \\ & + \frac{\phi_4}{|p_{2,i+1}-p_{2,i-1}||p_{1,j+1}-p_{1,j-1}|}|p_{2,i}-p_{2,i-1}||p_{1,j}-p_{1,j-1}|.\end{aligned} \tag{4-13}$$

For the coefficient interpolation, the application of this interpolation method is very simple since each one of the LTI model coefficients is associated with the operating points (described with the combinations of both scheduling variables) of their LTI model. In output-input interpolation, this method can be imagined that each interpolation weights α_n associated

to each LTI models have their own lookup table. All of the points in those tables (in the known combination of scheduling variable) will have values of zero, except at the scheduling variable combinations associated to the operating points of the LTI model. In this way, each interpolation weight at any point in the scheduling variable trajectories can be calculated by such as written in equation 4-13.

As this interpolation does not requires an optimization to assign the interpolation parameters, the only way to improve the result of this optimization is to increase the amount of the identified LTI models within the operating region.

4-2 PEM identification procedure

The local LPV identification approach utilize LTI models of the distillation column which are identified at different operating points. In those operating points we have to identify LTI model that represent the local dynamics of the system. The simplest way identify this LTI model is to conduct a *prediction error minimization* (PEM) based identification. This identification are based on the capability of a model to predict future values of signals as candidate representatives of the systems. The identified model are identified by minimizing the prediction error of the model output with respect to the modelled system (Ljung 1999).

The PEM identification is applied by using a so called identification cycle:

1. Design an input signal for the system which excite the dynamic of the system on the interested operating points, and collect a system input-output data set $\mathcal{D}_N = \{y(k), u(k)\}_{k=1}^{N}$.
2. Select a model order that are sufficient to capture the system dynamics.
3. Select a model structure $\mathcal{M}$ and its identification criterion, then estimate the process model and noise model.
4. Do residual analysis to verify whether the true system is contained within the estimated model $\mathcal{S} \in \mathcal{M}$. If the test fail, go back to Step 2.
5. Estimate the variances of process model and the noise model, and check if they are small. If not, repeat the cycle from Step 1.

Each of the steps of the cycled are explained in detail in the sequel. The concepts on each step of the identification cycles will be explained from the viewpoints of the distillation column.

4-2-1 Input design

Ideally we would like to excite the system on each frequency in order to obtain valuable information about the system dynamics. In reality, an input excitation is often restricted due to limitation of the system such as the operational costs and/or the safety regulations.

To choose an appropriate excitation signal, we have to know priori information about the plant. In the works of (Chien and Ogunnaike, 1992) and (Gaikwald and Rivera, 1994) an input signal design guidelines for the distillation column have been comprehensively studied. In those literatures, the concept of applying a *Pseudo Random Binary Noise* (PRBS) was used to generate the excitation signal. The PRBS is a two level signal which is generated using shift registers and Boolean algebra. This signal is characterized by two parameters: the number of shift registers n_r, and the switching time T_{sw}. The design guidelines to generate a rich PRBS are based on the information of dominant time constant of the system.

$$T_{sw} \leq \frac{2.6\tau^L_{\text{dom}}}{\alpha_s} \quad 2^{n_r} - 1 \geq \frac{2\pi\beta_s\tau^H_{\text{dom}}}{T_{sw}} \tag{4-14}$$

with τ^H_{dom} is the highest dominant time constant , τ^L_{dom} is the lowest dominant time constant, and α and β are integers chosen by the user (will be explained further in this section). At each operating point, these time constants can have different values characterizing the change in the dynamical properties of the system. The simplest way to approximate the time constant, is to calculate rise time of system step response for each input-output channel of the system. This time constant will also determine the minimum data point on the identification which is 5-10 largest time constant, this would equal around 1600 number of data point per model order. For the choice of the α_s and β_s, as noted in Chien and Ogunnaike (1992), for the high purity distillation column it is desirable to have a very fast closed loop settling time compared to the open loop settling time. Therefore it is usually chosen that β_s have slight larger value if compared to α_s (around 5 for β_s and 4 for α_s).

4-2-2 Model order choice

The next step is to choose the efficient model order with low complexity but being able to sufficiently well capture the dynamics of the column. When deciding the model order, an *Empirical Transfer Function Estimate* (ETFE) is usually applied to get an insight in the system dynamics. However, since we have access to the high fidelity model, we already have the information of the model order via the linearization of the nonlinear model. The LTI model derived from the linearization around a valid operating points has a 110th order transfer function. This large-scale LTI model has disadvantages that besides being difficult to accurately identify(large variance) it also rises computational problems for the interpolation. To avoid the latter problem, model reduction applied on each linearized LTI model shows that the linear model of order 4 can describe the dynamics of original linear model without any valuable loss (See Figure 4-8).

4-2-3 Model structure choice

An accurate model choice will results in unbias model estimate with

$$\bar{E}\{\varepsilon(k,\theta)\} = 0, \tag{4-15}$$

and variance of

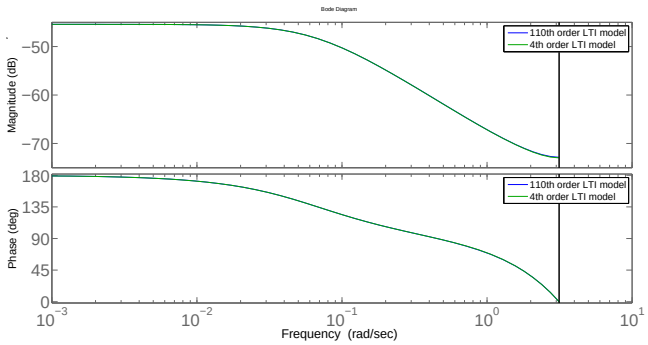


Figure 4-8: Frequency response comparison between 4th and 110th order LTI model.

$$\bar{E}\{\varepsilon^2(k,\theta)\} = \sigma_e^2. \tag{4-16}$$

with $\varepsilon(k,\theta) = \hat{y}(k|k-1;\theta)$ is one step ahead prediction error, and σ_e^2 is variance of the white noise. This statement means that the powers of prediction error of an unbiased model, will be equal to the variance of the zero mean white noise.

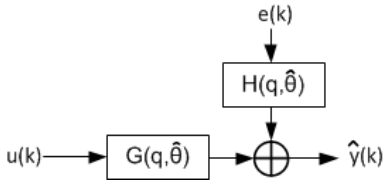


Figure 4-9: PEM model structures.

To select the appropriate model structure, it is better to understand what type of noise is affecting the system. In Section 2-3-1 we have decided that we will only consider measurement noise for this system. This consideration lead us to an *output error* (OE) model class for local LTI models of the plant

$$G(q,\hat{\theta}) = \frac{B(q,\hat{\theta})}{F(q,\hat{\theta})}; \qquad H(q,\hat{\theta}) = 1, \tag{4-17}$$

with $G(q,\hat{\theta})$ is the estimated process model, and $H(q,\hat{\theta})$ is the noise model. For OE model structure, the predictor output ($\hat{y}(k,\theta)$) is nonlinear in θ. Therefore, the estimated parameter defined as

$$\hat{\theta}_N = \arg\min_{\theta} \mathcal{V}_N(\theta, \mathcal{D}_N), \tag{4-18}$$

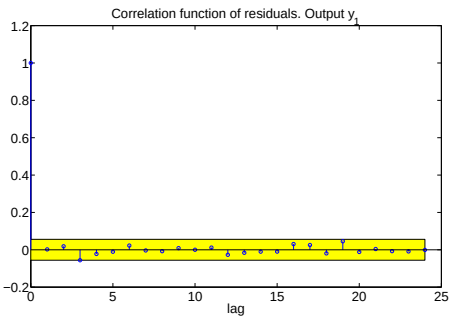
with identification criterion

$$\mathcal{V}_N(\theta, \mathcal{D}_N) = \frac{1}{N}\sum_{k=1}^{N} \varepsilon^2(k, \theta), \tag{4-19}$$

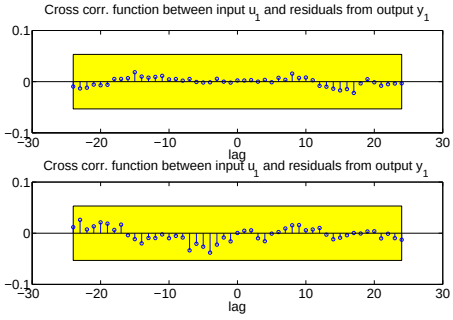
can not be determined analytically, and its solution have to be computed iteratively with risk of finding a local minimum.

4-2-4 Validation

At the end of each identification cycle, the model have to be validated in term of its bias as well as its variance with respect to the data-generating system. To conclude a model is unbiased with respect to underlying system, the identified model have to undergo a residual test to test the hypotheses whether the true system is contained inside the model $\mathcal{S} \in \mathcal{M}$. Residual analysis consists of two tests, the whiteness test and the independence test. According to the whiteness test criterion, a good model has the residual autocorrelation function inside the confidence interval of the corresponding estimates, indicating that the residuals are uncorrelated. For the independence test criteria, a good model has cross correlation of the residuals uncorrelated with past inputs (See Figure 4-10).



(a) Auto Correlation of the first output(Bottom Quality)



(b) Cross Correlation between first output with both inputs (Vapor and Liquid Flows)

Figure 4-10: Residual Test example.

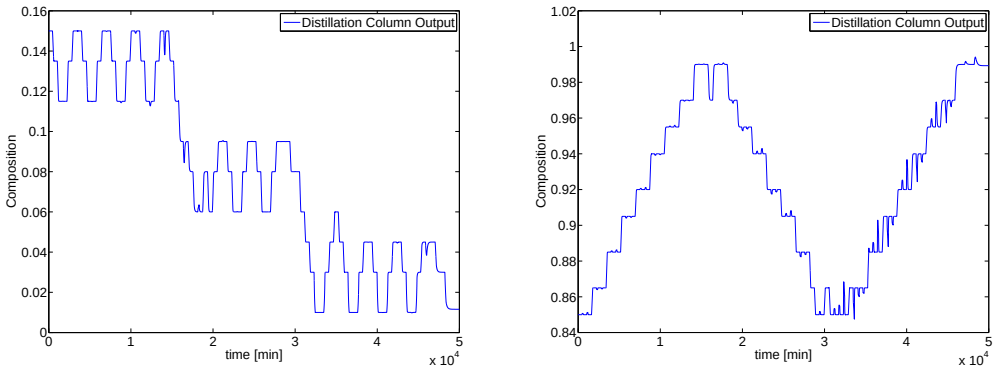
After the identified model have proven to be unbiased, the next task is to determine its variance with respect to the original system. This validation is done in the frequency domain, by calculating the standard deviation of the estimated model. At each frequency, the modeling error is very likely to be small if the standard deviation is small

$$\frac{\sqrt{cov(G(e^{j\omega}, \hat{\theta}_N))}}{|G(e^{j\omega}, \hat{\theta}_N)|} < 0.1 \quad \forall \quad \omega \in [0\ 1]. \tag{4-20}$$

with 0.1 means that the model are identified with 99%-confidence regions. If the standard deviation is too big, the excitation signal need to be revised. The variance of the identified model can be reduced by increasing the number of data N or increasing the power spectrum $(\Phi_u\omega)$ of the frequencies where the variance was too large.

4-3 Result and analysis

In this section we will show the result that we have found in the application of the local LPV identification approach on the data set obtained from the high fidelity model. For these approach, we choose the output of the distillation column, i.e., top and bottom product composition, as the scheduling variables for the LPV system. The choice of these scheduling variables is highly motivated since the top and bottom product composition directly represents the operating points of the system. Hence, it automatically characterize the operation of the system.



(a) First scheduling variable **(b)** Second scheduling variable

Figure 4-11: Scheduling variable for local LPV identification approaches.

Afterwards, there are some items that needs to be decided in the context of local LPV identification approach. Those items are included in these question:

- How many LTI model do we need? At which point of the system operating regime should we do the identification?
- What interpolation scheme should we choose for this system?
- What interpolation method should we choose for this system?

Unfortunately, there are no guidance on answering these three items. Intuitively we can expect an answer for this question based on the available information of the distillation column itself. For instance, due to the fact that we choose the output of the distillation column as the scheduling variables, we expect better result in the output interpolation scheme. Since we have the access to the nonlinear model of the distillation column, we will do preliminary experiments to analyze the effect of different interpolation schemes, method, and numbers of the LTI needed for the interpolation (See Section 4-3-1). We will use the results of this study to establish a better insight to the effect of these choices and assess which is the most beneficial approach among the possible combinations in the local LPV identification of the distillation column. After choosing the best candidate among schemes, methods, and needed numbers of LTI models, we can proceed with the interpolation experiment by using LTI model obtained via PEM identification.

For the whole identification procedure, the local identification approach requires 2 until 3 types of data set. The first one is the identification data sets ($\mathcal{D}_N$) from each desired operating points of the system. This data set contain local (frozen) dynamic information of the system, and are used to obtain the local PEM model of the system, and have minimum 6400 data point (See Section 4-2). The second one is validation data sets ($\mathcal{D}_V$) which are used to validate the quality of the obtained LPV model with respect to varying scheduling variable. This validation data set contain 20000 data point. The third one is parameter optimization data set for output and input interpolation scheme ($\mathcal{D}_O$), which should contain rich information on the dynamics of the system within its operating regimes (50000 data point) (See Figure 2-7). This data sets are not necessary for the RBF and bilinear based interpolation, but required for the polynomial based interpolation or to optimize the weigh in the RBF based interpolation. The data set can be seen in appendix B

The result of each experiment are presented in terms of *Means-Squared Error* (MSE) and *Best Fit Ratio* (BFR) with respect to validation data set $\mathcal{D}_{\text{val}} = \{y(k), u(k), p(k)\}_{k=1}^{N}$. The MSE is defined as:

$$\text{MSE} = \text{mean}((Y - \hat{Y})^2) \tag{4-21}$$

while BFR is defined as:

$$\text{BFR} = 100\%.\max\left(1 - \frac{\| Y - \hat{Y} \|_2}{\| Y - \bar{Y} \|_2}\right) \tag{4-22}$$

with $Y = \begin{bmatrix} y(1) & y(2) & \dots & y(N) \end{bmatrix}^\top$, is the output of the original system, and $\hat{Y} = \begin{bmatrix} \hat{y}(1) & \hat{y}(2) & \dots & \hat{y}(N) \end{bmatrix}^\top$ is the output of the LPV model.

4-3-1 Analysis of local LPV identification approach for the high-purity distillation column

Most of the steps to conduct local LPV identification approach are done in an ad-hoc manner. Hence, it would be beneficial for the identification studies to explore the effect of each combinations of interpolation schemes and methods for the distillation column. To analyze each of the interpolation combinations, we will conduct preliminary test by using LTI models obtained via linearization of the nonlinear model instead of LTI models obtained from PEM identification. This test can be done because we have the access to the nonlinear model where we can linearize it at will. We can produce as many LTI model of the system necessary for the interpolation, even to produce large amount of LTI model identified at various point within the system operating regime that will guarantee us to capture the nonlinear dynamics of the system. However, having to many LTI model for the interpolation is unlikely to be done since the effort and the cost on doing the PEM identification will increase, and the complexity of the interpolation may increase as well.

This preliminary experiment will be done by comparing the combinations of scheme and method with respect to different numbers of identified LTI models. The location of these LTI models are chosen by equidistant gridding of the operating regime of the system which

are 0.01-0.15 on the bottom composition and 0.85-0.99 on the top composition (See Section 2-3-1). The numbers of operating regime gridding that are considered in the experiments are:

- 5x5 LTI model,
- 8x8 LTI model,
- 15x15 LTI model, and
- 29x29 LTI model.

By doing this experiments we will be able to determine the optimal combinations of interpolation scheme, methods, and needed amount of local LTI model for the distillation column.

RBF based interpolation

The RBF based interpolation have flexibility to assign their parameters with or without an optimization. If we decide to do the optimization. This allows to incorporate fitting of the interpolation function to the intersample local behaviors of the system, but also to refine the model response during transients between the local dynamic descriptions. On the table 4-2 we can see a comparison example of RBF interpolation on the output interpolation scheme with and without an optimization.

Table 4-2: Comparison of simulated LPV model output fit ratio under optimized and manually assigned RBF interpolation weights

	With Optimization		Without Optimization	
	MSE	BFR	MSE	BFR
Training Data set	$1.638.10^{-5}$	92.29%	$1.723.10^{-5}$	91.79%
Validation Data set	$5.047.10^{-6}$	94.88%	$5.723.10^{-6}$	93.71%

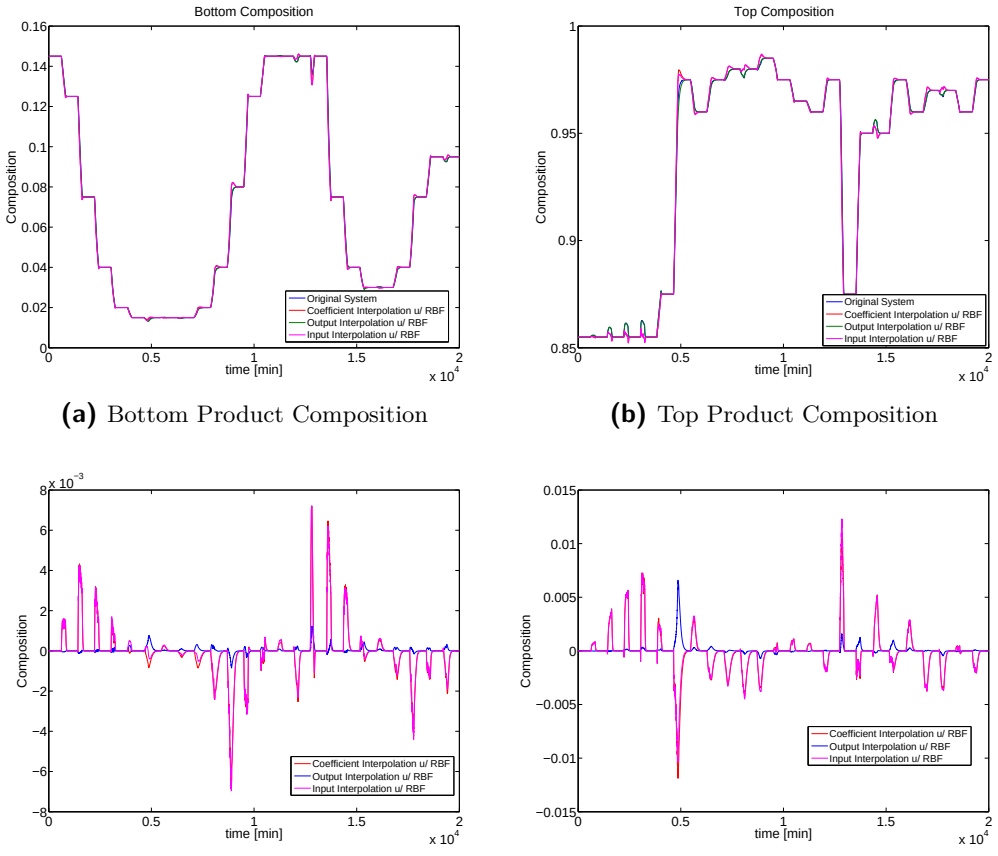
The table shows that parameters optimization enhance the result of the manually selected RBF parameters. Since we have the access to a high fidelity distillation column model, conducting experiment to obtain the data set used for the optimization is not an issues for us. In practice, such a data is also often available, due to previous operation of the column which can be used for history matching. On the further experiment, we will only consider RBF results obtained via parameter optimization.

To asses the three interpolation schemes under RBF based interpolation method, we have done multiple experiments using various number of LTI models such as explained in Section 4-3-1. The result can be seen in Table 4-3:

This table shows that the input interpolation have a worse result if compared to the output interpolation scheme. This is understandable since we choose the output of the system as the scheduling variable. The result of the coefficient interpolation in terms of these error measures is slightly worse if compared to the output interpolation. Another thing to keep in mind is that the complexity and the computational load of the optimization will increase if we have more LTI models involved in the interpolation scheme.

Table 4-3: comparison of simulated LPV model output fit ratio under RBF interpolation method

	Coefficient Interpolation		Output Interpolation		Input Interpolation	
	MSE	BFR	MSE	BFR	MSE	BFR
29x29 LTI models	$2.350.10^{-6}$	96.99%	$1.487.10^{-7}$	99.31%	$2.328.10^{-6}$	96.93%
15x15 LTI models	$2.660.10^{-6}$	96.66%	$4.240.10^{-7}$	98.66%	$2.666.10^{-6}$	96.66%
8x8 LTI models	$4.405.10^{-6}$	95.31%	$4.169.10^{-6}$	96.01%	$6.169.10^{-6}$	93.38%
5x5 LTI models	$1.998.10^{-5}$	91.07%	$5.047.10^{-6}$	94.88%	$2.719.10^{-5}$	89.48%



(a) Bottom Product Composition

(b) Top Product Composition

(c) Error on Bottom Product Composition

(d) Error on Top Product Composition

Figure 4-12: Comparison of the true and the simulated LPV model output in terms of the output error under RBF interpolation.

Polynomial interpolation

In the investigation of the polynomial interpolation method we will not consider the input interpolation scheme anymore. The main reason is that from the previous results we can see that the input interpolation scheme has inferior results if compared to the output interpolation

scheme. As both schemes directly utilize LTI model output, with the previous RBF interpolation result, we can expect the input interpolation to produce a worse results for capturing the distillation column dynamics. The other reason, is that this polynomial interpolation require an optimization to select the polynomial function parameter θ. The amount of parameters that need to be estimated are based on the order of the polynomial, and this amount will be much larger than the previous RBF interpolation method (See Section 4-1-2). The nonlinear optimization algorithm written in Matlab cannot efficiently solve such a large-scale problem. In the output interpolation scheme however, solving the output interpolation is nothing more than solving a LS problem which results in the unique solutions and less complex problem if compared to the nonlinear optimization. Even for a LS problem, with many LTI models involved (15x15 and 29x29 gridding), the memory restriction in Matlab restrict us from constructing such a large matrix. This way we only able to use lower order polynomials, which are second order polynomial for 29x29 numbers of LTI models, and fourth order polynomials for 15x15 numbers of LTI models. As for the coefficient interpolation there is no restrictive issue encountered during this experiment.

Table 4-4: Comparison of simulated LPV model output fit ratio under polynomial interpolation method

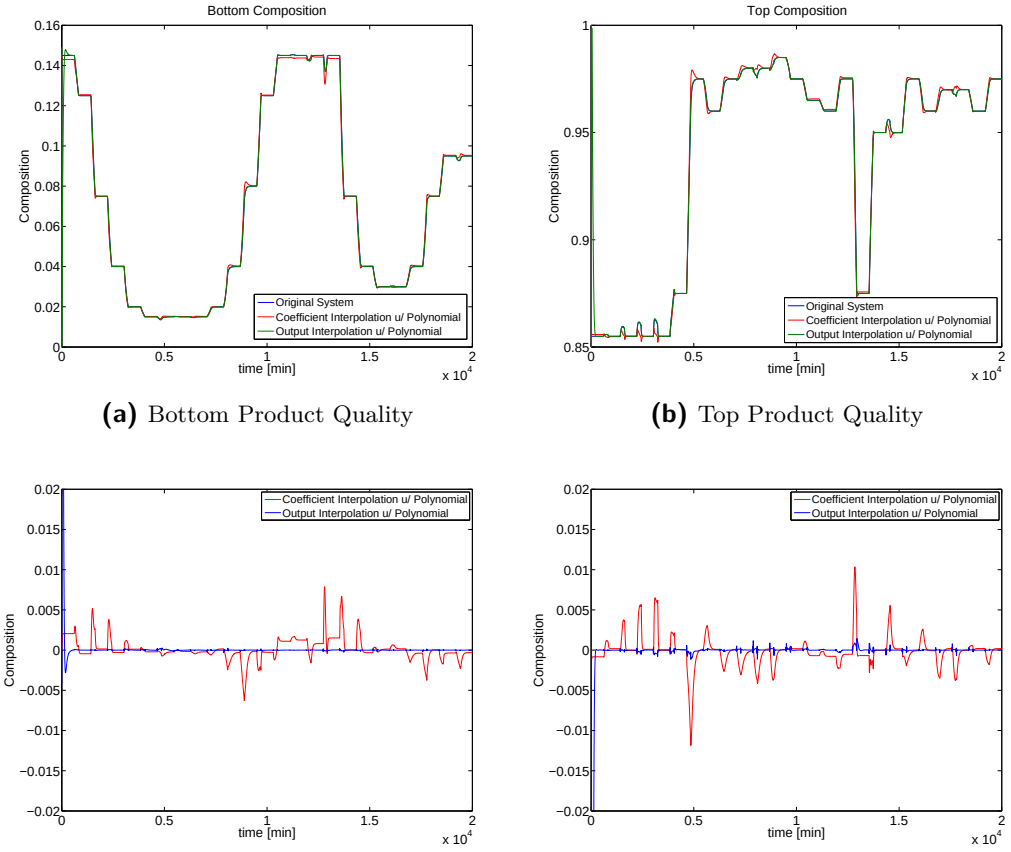
	Coefficient Interpolation		Output Interpolation	
	MSE	BFR	MSE	BFR
29x29 LTI models	$2.505.10^{-6}$	96.80%	$1.480.10^{-5}$	94.14%
15x15 LTI models	$2.349.10^{-6}$	96.94%	$1.676.10^{-5}$	93.77%
8x8 LTI models	$3.821.10^{-6}$	95.99%	$1.385.10^{-5}$	94.23%
5x5 LTI models	$1.683.10^{-5}$	91.99%	$3.359.10^{-5}$	91.20%

If we compare the previous RBF based interpolation result (Table 4-3) with the polynomial based interpolation result (Table 4-4), we can see two main differences. The first one is a slight improvement in the accuracy of the LPV model derived using coefficient interpolation scheme especially on the interpolation which involve less LTI models. The second difference we can see is that, the result of the output interpolation scheme appears to be worse if compared to the RBF based interpolation result.

The improvement on the coefficient interpolation can be explained because by using polynomial based method, we fit a surface of polynomial function in the measured points of the LTI model parameters (model coefficient, and input-output trimming values). In this manner, new model parameters value (parameters which are not associated with any identified LTI models), are calculated by not just the surrounding model parameters value, but by the trends of the parameters in the whole operating regimes of the system. This is why the improvement is even more significant when less LTI model parameters are involved for the interpolation. If the model parameters are well measured over the operating regimes (many LTI models are used for the interpolation) the polynomial interpolation will not further improves the accuracy of the model.

On the other hand, the output interpolation suffers loss of accuracy in this interpolation method. This happens because when we use a high order polynomial as the interpolation

weight. The surface of the weight can become a bit bent at the edge of the operating regime of the system. This interpolation however, have a very good results when the LPV model is simulated far from the edge of the operating regime (can be further seen in the Figure 4-13).



(a) Bottom Product Quality

(b) Top Product Quality

(c) Error on Bottom Product Composition

(d) Error on Top Product Composition

Figure 4-13: Comparison of the true and the simulated LPV model output in terms of the output error under polynomial interpolation.

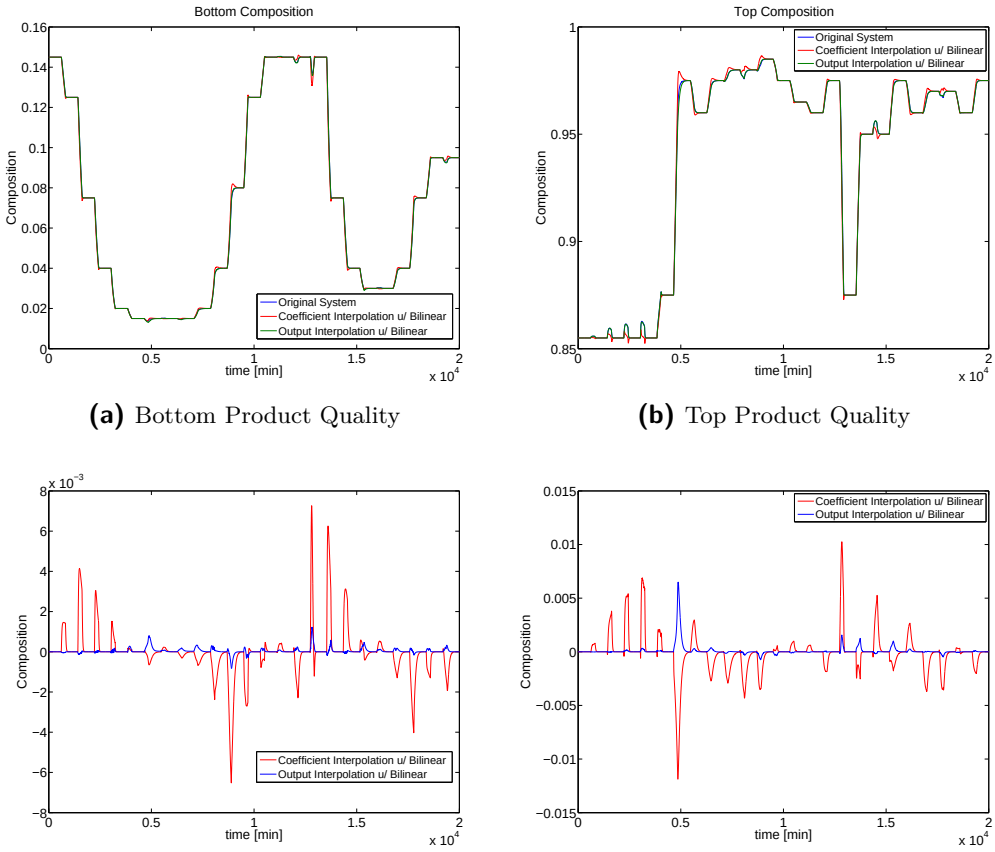
Bilinear interpolation

As previously explained in Section 4-1-2, bilinear interpolation is a straightforward interpolation method where we does not need to assign any parameters for the interpolation weight. Hence, the result of bilinear interpolation only depends on the amounts of LTI models or the location of the operating points of the LTI models. The comparison between the output interpolation and coefficient interpolation schemes, can be seen in Table 4-5:

As an additional comparison, we also investigate the effect of using massive amount of LTI models using 61x61 gridding of the operating region. With this much LTI models used to describe the local dynamics of the system, we can get an insight to the near maximum result that we can achieve when apply local LPV identification approach. The result however, does not express a significance increase in accuracy if compared to the 29x29 gridding.

Table 4-5: Comparison of interpolation scheme under bilinear interpolation method

	Coefficient Interpolation		Output Interpolation	
	MSE	BFR	MSE	BFR
61x61 LTI models	$2.251.10^{-6}$	97.00%	$1.487.10^{-7}$	99.30%
29x29 LTI models	$2.249.10^{-6}$	96.99%	$1.487.10^{-7}$	99.30%
15x15 LTI models	$2.585.10^{-6}$	96.71%	$4.313.10^{-7}$	98.65%
8x8 LTI models	$5.815.10^{-6}$	95.07%	$4.599.10^{-6}$	95.80%
5x5 LTI models	$2.972.10^{-5}$	89.22%	$3.017.10^{-5}$	89.23%



(a) Bottom Product Quality

(b) Top Product Quality

(c) Error on Bottom Product Composition

(d) Error on Top Product Composition

Figure 4-14: Comparison of the true and the simulated LPV model output in terms of the output error under bilinear interpolation.

Frequency domain comparison

Furthermore, beside the simulated output, we also compare the frequency response of the LPV models with respect to variation of scheduling variable (different operating points within the

operating regimes of the system) (See Appendix A). The comparison of the frequency response are useful especially from the control synthesis point of view. From the figures, we can see that the LPV model obtained via coefficient interpolation have smoother transition between the operating points of the system, even on the scheduling variable values which are not associated with any LTI model operating points. In contrary we can see some additional resonance peak which abruptly occurred in the LPV model obtained from output interpolation. This happen due to the fact that we optimize the parameter of the interpolation weight such that the output fit the optimization data set. An even worse effect from this fitting are shown in the polynomial based interpolation on the output interpolation schemes. This behavior will certainly give difficulties for the control design of the distillation column.

Conclusion on the preliminary experiment

Summing up the results from all the previously discussed findings, we can arrive at several conclusions that can help answer earlier questions about conducting local LPV identification in the distillation column. Based on the result, we could see that the 15x15 LTI models based interpolations are already enough to capture the system dynamics. Using more LTI models for the interpolation will only produce an insignificant increase in accuracy, but on the other hand, it significantly increase the computational load of the interpolation. The time needed to solve the parameters optimization problem on both RBF and polynomial interpolation is significantly longer when we use more LTI models than 15x15 gridding. Using less LTI models (8x8 or 5x5 gridding) will lower the fit result for each combination of interpolation methods and schemes by around 2%, although this would mean that we could reduce the cost of the identification result if we have to do PEM identification on real system as well as to reduce the computational load for the parameters optimization.

To decide the best combination of interpolation schemes and methods, we can see from two sides which is the output response and the frequency response of the LPV models. From the output response, the LPV models output derived from different combinations of interpolation schemes and methods does not show significant difference from one to the other combinations. However, the output interpolation using polynomial interpolation are clearly bellow the other combinations of interpolation methods and schemes. In the frequency response comparison however, it is clearly seen that LPV model based on output interpolation schemes are unlikely to be chosen if we would want to synthesis a controller for the distillation column.

Based on this result we decide to apply coefficient based interpolation schemes with polynomial interpolation method on the distillation column setup. The further step is to obtain local model of the system by using PEM identification method.

4-3-2 PEM identification on high-purity distillation column

The PEM identification routines has been carried out in all 15x15 chosen operating point. The elements of identification routine including the input excitation design, model structure, model class choice, validation, etc have been explained in Section 4-2. The routines are done under three different noise conditions such as low noise condition with 40dB *Signal to noise Ratio* (SNR), medium noise condition of 25dB SNR, and severe noise condition of 10dB SNR.

The SNR is a measure that compares the power of the desired signal (output signal $y(k)$) to the power of background noise ($e(k)$), which is formulated as:

$$\mathrm{SNR}_{\mathrm{dB}} = 10 \log_{10} \frac{P_{y(k)}}{P_{e(k)}}, \tag{4-23}$$

with the signal power $P_{x(k)}$ is defined as

$$P_{x(k)} = \frac{1}{N} \sum_{k=0}^{N-1} x^2(k). \tag{4-24}$$

The noise considered in this experiment is a measurement white noise, and its variances is scaled with respect to magnitude of the distillation column output to produce the specific SNR mentioned above.

PEM identification result

Based on the identification result, we found that the estimated model coefficients of the identified OE LTI model is not the same with the model coefficient of the LTI model achieved via linearization. This is not a problem, since we are sure that the true model is included via the residual analysis, and the estimated LTI model output that have a high fit ratio on the validation data set. The difference on the model coefficient, is just a matter of different realization of the system. Upon calculating the variance, we face a problem because of the low static gain of the system. This results in the calculation of the estimate of the covariance matrix $\hat{P}_\theta$ became numerically ill conditioned. We decide to do 100 loop of *monte carlo* run of PEM identification at each operating points (with same excitation signal but different noise generated in each run), and then estimate the variance from the calculation of the standard deviation of each OE model coefficient directly from the result of the monte carlo run. An example of identified OE model coefficient and its standard deviation in an arbitrary chosen operating points can be found in Table 4-6 and Table 4-7

Table 4-6: Model coefficient of the identified OE model at an operating point in low noise condition

	Denumerator
mean	[1.000 -2.395 1.715 -0.191 -0.126]
STD	[0 0.888 1.366 0.913 0.361]
	Numerator 1
mean	1.10^{-3} * [-0.0014 0.3761 -0.5755 0.1672 0.0409]
STD	1.10^{-5}*[4.817 10.504 35.764 35.524 19.320]
	Numerator 2
mean	1.10^{-3} * [-0.0037 -0.3847 0.6010 -0.1792 -0.0413]
STD	1.10^{-5}*[3.092 2.342 35.779 31.693 15.944]

Table 4-7: Model coefficient of the identified OE model at an operating point in severe noise condition

	Denumerator
mean	[1.000 -0.444 0.090 -0.373 -0.128]
STD	[0 1.011 1.110 0.765 0.608]
	Numerator 1
mean	1.10^{-3} * [0.0067 0.4114 0.1298 0.1784 -0.0345]
STD	$1..10^{-4}$*[1.717 3.359 6.203 4.873 2.80]
	Numerator 2
mean	1.10^{-3} * [0.0077 -0.4477 -0.1374 -0.2181 0.0687]
STD	$1..10^{-4}$*[1.612 2.813 5.142 4.258 3.095]

In Table 4-6 and 4-7 , it can be seen that the variance of the model parameters is large (seen by the large standard deviation) , especially under severe noise conditions (SNR=10db). This happens because the PEM identification in the OE settings have a chance that each identification will end in local minima.

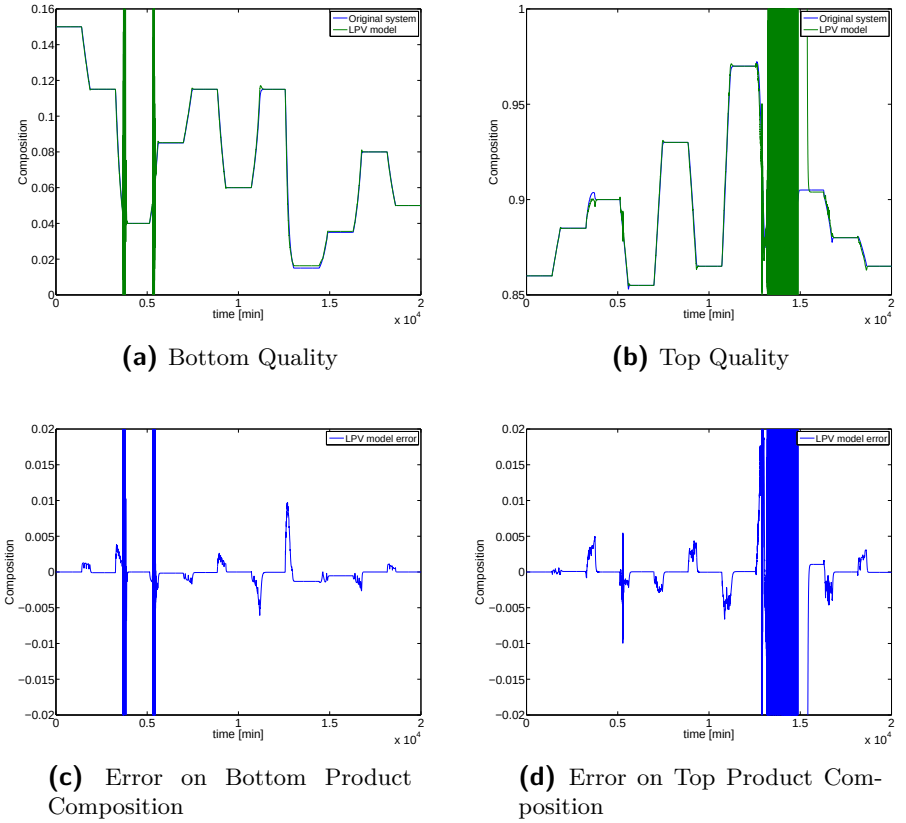
The best thing that we can do to select the OE model that will be used for the interpolation, is to select the highest fit ratio (BFR) (Eq. 4-22) of its estimated output with respect to the original output (in the identification dataset). Additionally, since on each operating point (of the 15x15 gridding) we have LTI model of obtained via linearization of the nonlinear model, we could also calculate the fit ratio of the OE model frequency response with respect to the LTI model frequency response. Although this frequency response comparison might not be available if the identification is done in the real distillation column plant, the BFR of the frequency response can help select the best OE model from the monte carlo run.

Another thing that is found in the identification process, is that some of the identified OE model are seen to have multiple pole near the unit disc. Due to the inability to correctly calculate the covariance matrix, the only thing we can do is to take precaution of the variance of the identified OE models when we are doing simulation with the LPV model. The precaution step that can be done, is to limit the aggressiveness of the input signal such that this signal will not evoke the high frequency mode of the OE models used in the LPV system.

4-3-3 Application of local LPV identification approach using identified OE model of high-purity distillation column

After we have identified OE models on each of points in the 15x15 gridding of the system operating regimes, we further interpolate those local model in the chosen combination of interpolation schemes and methods which is the coefficient interpolation and polynomial based interpolation. The simulation on the LPV models shows stability issues in the output response, which can be seen in Figure 4-15.

At first we are still not sure why does the stability issue came out even when all the identified OE model are a stable model. We then decide to further investigate wether this problem also occurred if we are using different interpolation method for the coefficient interpolation



(a) Bottom Quality

(b) Top Quality

(c) Error on Bottom Product Composition

(d) Error on Top Product Composition

Figure 4-15: Stability Issue in coefficient interpolation.

scheme. From the further experiment, it turns out that the stability issue also happen by either using RBF and bilinear based interpolation method. The output error comparison for the coefficient interpolation scheme can be seen in Table 4-8.

Table 4-8: Comparison of simulated LPV model output fit ratio under coefficient interpolation scheme

	RBF		Polynomial		Bilinear	
	MSE	BFR	MSE	BFR	MSE	BFR
Low Noise	$7.73.10^{-4}$	45.96%	$8.173.10^{-5}$	77.019%	$9.945.10^{-5}$	71.75%
Medium Noise	$5.617.10^{+4}$	48.36%	∞	0%	$1.144.10^{-2}$	30.64%
severe noise	∞	0%	$3.156.10^{-4}$	49.25%	∞	0%

This result is very bad if we compare it with the LPV model output response obtained in the preliminary experiment (Section 4-3-1). Although in this section, the OE model that we used for the interpolation is basically not as perfect as the LTI model achieved via linearization of the nonlinear model. We cannot really answer why does the stability issue occurs in different LPV model which derived from different combination of OE model set (identified at different noise conditions) and interpolation method. However, it appears that even if the model

coefficients used for the coefficient interpolation are taken from a stable OE model, when the LPV model travel through the operating region of the system, the combination of LPV model coefficient might result in unstable pole. This unstable pole is the one that cause stability issues. This is sheer consequences since we do the interpolation in the frequency domain (Tóth 2007).

To better investigate wether the modeling error in the OE model effect the output interpolation scheme, we also conduct the interpolation using the output interpolation scheme also with various interpolation method. The comparison of the simulation result can be seen in Table 4-9.

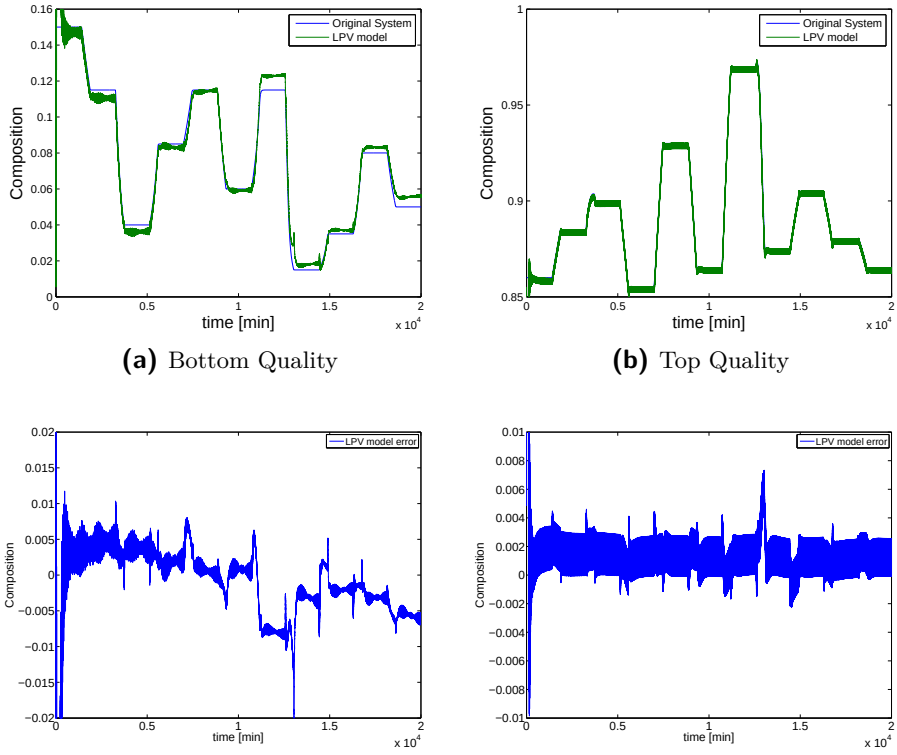
Table 4-9: Comparison of simulated LPV model output fit ratio under output interpolation scheme

	RBF		Polynomial		Bilinear	
	MSE	BFR	MSE	BFR	MSE	BFR
Low Noise	$4.219.10^{-5}$	83.23%	$4.193.10^{-5}$	45.18%	$3.993.10^{-5}$	83.71%
Medium Noise	$2.648.10^{-5}$	85.89%	$7.877.10^{-4}$	37.39%	$2.550.10^{-5}$	86.24%
severe noise	$6.413.10^{-4}$	41.07%	$1.315.10^{-5}$	49.13%	$6.407.10^{-4}$	41.51%

When the LPV model is derived with the output interpolation scheme, we do not see anymore severe stability issue on the LPV model. However the accuracy of the LPV model obtained via interpolation of OE model identified on severe noise conditions are still very low. This is because at that noise condition the identified OE models have much lower accuracy if compared to the other noise conditions. The sensitive of the low accuracy OE model is inherited to the LPV model that are obtained by interpolating those models. Another thing that can be seen in this table, is that LPV model obtained via polynomial based interpolation method, have low fit ratio for all three noise conditions. The result of polynomial based interpolation method can be seen from Figure 4-16.

The previous issues about the bad fit ratio when the points is located on the edge of the operating regime are still there (See Section 4-3-1). This will punish the result of the output interpolation scheme when using polynomial interpolation. Moreover, from that figure we can also see oscillations behavior in the output simulation result. This is still the same inherited behavior from the OE models. However, for the polynomial function, it is more clearly seen since the constraint used for interpolation weights optimization are not as restrictive as the RBF interpolation. The interpolation method will allow more models to be active at many points of the operating regime. This will surely raise the chance to have active models that have poles located in the imaginary axis.

Additionally, based on the PEM identification result, we know that at several operating points, the identified OE models have multiple poles near the unit circle, the OE models in that points are sensitive and will only be able to give valid system response in a very close vicinity of the corresponding operating points. When the input signal differs a bit from the operating condition, those LTI models may produce output with huge magnitude. This problem will lead to steady state output offset and poor transient behavior. Another problem that are also found is that if the local model have pole located on imaginary axis, oscillation behavior may also seen in the identified LPV system.



(a) Bottom Quality

(b) Top Quality

(c) Error on Bottom Product Composition

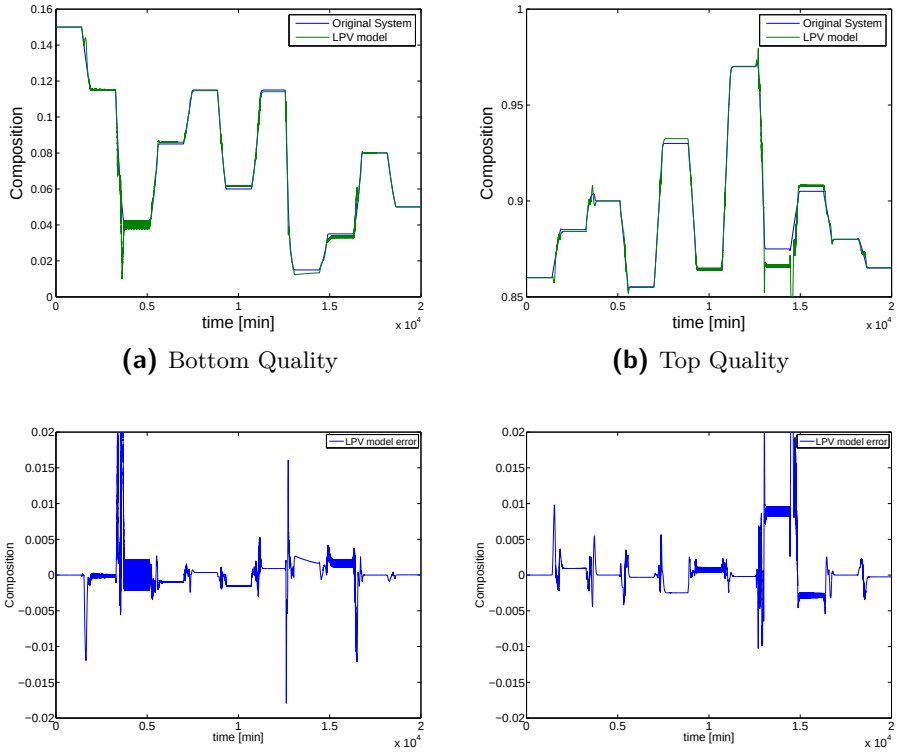
(d) Error on Top Product Composition

Figure 4-16: Comparison of the true and the simulated LPV model output in terms of the output error under polynomial interpolation.

4-4 Summary on the local LPV identification approach

The local LPV identification approach utilize LTI models of the studied system. In order to obtain those models this thesis incorporate PEM identification theories to identify the LTI models. A useful design guidelines are also found to provide more realistic way to excite the distillation column. The identification on the PEM identification faces several problems due to the low gain of the distillation column. The covariance matrix became numerically bad and singular, hence it is unreliable to be used to calculate the variance of the estimated parameters.

In the end, the quality of local LPV identification approach really depends on the identified LTI model quality. On the result section, where the LPV identification are done using LTI model achieved via linearization, we see very good result when we use LTI models derived from the linearization of the original system. However, due to LTI model identification error, the result of the LPV identification using LTI models achieved via PEM identification shows low result. The severe noise and system with high nonlinearities may give problem for this linear identification. The result of interpolation with low quality model may induce an LPV system with output offset, oscillation problems, as well as stability problem. The decision to conduct a local LPV identification approach on a studied system have to be considered



(a) Bottom Quality

(b) Top Quality

(c) Error on Bottom Product Composition

(d) Error on Top Product Composition

Figure 4-17: Comparison of the true and the simulated LPV model output in terms of the output error under RBF interpolation.

carefully as the interpolation of LTI models with huge estimation errors might results in various described problems.

Chapter 5

Global LPV identification of High-Purity Distillation Column

In this chapter, global LPV identification of the high purity distillation column using an LPV-ARX (*auto-regressive model with exogenous input*) model structure is studied. This model structure can be seen as the extension of the LTI-ARX model structure, where the model coefficients are estimated via analytical solution of a *least-squares* problem. However, in the LPV setting, the model coefficients are no longer constants, but functions of the scheduling variable. Due to this fundamental difference, the coefficient estimation in the LPV case express the nonlinearities of the systems in terms of a function dependent on scheduling variable. The commonly applied identification approach will need a best possible basis selection to expand the system nonlinearities. Inaccurate selection will lead to structural bias, while over parametrization will result in a variance increase of the estimate. Moreover, if there is no prior information about the nonlinearities of the system, the selection of the optimal function dependencies is not an easy task. Instead of pondering into the selection of the functional dependencies, the global LPV identification approach presented in this thesis adopt an alternative approach based on "non-parametric" estimation of the system nonlinearities. This approach cames from adopting a certain machine learning approach called *support vector machine* (SVM), in particularly the least-square formulation of SVM, which is named LS-SVM (*Least Square Support Vector Machine*). In this "non-parametric" approach, instead of explicitly stating a parametrization of the LPV model coefficients can be calculated directly on each scheduling variable. The methodology of the LPV LS-SVM identification can handle the large dimensional aspect of the high purity distillation column, and the identification results shows that this approach appears to be an efficient data-driven identification based modeling solution for high-purity distillation columns.

5-1 MIMO LPV-ARX formulation

The LPV-ARX model structure is an extension of the polynomial LTI model structure with an ARX noise model.

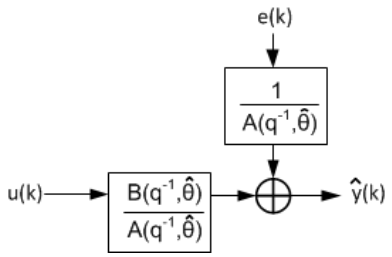


Figure 5-1: LTI-ARX model structure

The ARX model structure is often motivated due to the linearity in the parameter properties of this model structure. This will allow an analytical expression for the optimal model parameter θ that minimize the identification criterion (See Section 4-2-3). In these model, the deterministic part of the data-generating system is written in an input-output representation form:

$$y(k) = \sum_{i=1}^{n_a} A_i q^{-i} Y(k) + \sum_{j=0}^{n_b} B_j q^{-j} u(k) + e(k) \tag{5-1}$$

with output vector $y(k)$, input vector $u(k)$, and coefficient matrices A_i and B_j:

$$y(k) = \begin{bmatrix} y_1(k) \\ \vdots \\ y_{n_y}(k) \end{bmatrix}, \quad u(k) = \begin{bmatrix} u_1(k) \\ \vdots \\ u_{n_u}(k) \end{bmatrix}, \tag{5-2}$$

$$A_i = \begin{bmatrix} a_{i,1,1} & \dots & a_{i,1,n_y} \\ \vdots & \ddots & \vdots \\ a_{i,n_y,1} & \dots & a_{i,n_y,n_y} \end{bmatrix}, \qquad B_j = \begin{bmatrix} b_{j,1,1} & \dots & b_{j,1,n_u} \\ \vdots & \ddots & \vdots \\ b_{j,n_y,1} & \dots & b_{j,n_y,n_u} \end{bmatrix}. \tag{5-3}$$

Here $k \in \mathbb{Z}$ denotes the discrete time and $e(k) = [e_1(k) \quad e_{n_y}(k)]^\top$ is a white stochastic noise process. In this formulation, each one of the output is influenced by previous values of all the system outputs $[y_1(k) \quad \dots \quad y_{n_y}(k)]^\top$. The identification of the original dynamics in this form is often seen as the estimation of a non parsimonious representation resulting in higher variance and more computational load if compared to the simplification of the same MIMO (*Multi-Input-Multi-Output*) model via the assumption of a common denominator form. In this manner, the system is transformed into a different representation, such that the considered model coefficient A became:

$$A_i = \begin{bmatrix} a_{i,1} & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & a_{i,n_y} \end{bmatrix}. \tag{5-4}$$

Through this formulation, each system output channel will no longer has a direct influence on the other system outputs. The coupled output dynamics are now implicitly represented

in the newly transformed B_j matrix. In this way, identification of a MIMO system can be done separately on each output channels via the estimation of multiple *Multi-Input-Single-Output* (MISO) systems. Although this will result in a slight drawback of higher order for the channels. In the same manner, when dealing with a MIMO system, LPV-ARX model structure can follow a similar reformulation:

$$y(k) = \sum_{i=1}^{n_a} A_i(p(k))q^{-i}y(k) + \sum_{j=0}^{n_b} B_j(p(k))q^{-j}U(k) + e(k) \tag{5-5}$$

with

$$A_i(p(k)) = \begin{bmatrix} a_{i,1}(p(k)) & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & a_{i,n_y}(p(k)) \end{bmatrix} \quad B_j(p(k)) = \begin{bmatrix} b_{j,1,1}(p(k)) & \dots & b_{j,1,n_u}(p(k)) \\ \vdots & \ddots & \vdots \\ b_{j,n_y,1}(p(k)) & \dots & b_{j,n_y,n_u}(p(k)) \end{bmatrix} \tag{5-6}$$

where $p : \mathbb{Z} \to \mathbb{P}$ denotes the scheduling variable with range $\mathbb{P} \subseteq \mathbb{R}^{n_p}$. As said previously, model coefficients $A_i(p(k))$ and $B_j(p(k))$, are functions which vary through the instantaneous value of $p(k)$. In order to identify the model coefficient, the system structure in Eq. 5-5 need to be rewritten into a linear regression form. For each output

$$y_n(k) = \sum_{i=1}^{n_a} a_{i,n}(p(k))q^{-i}y_n(k) + \sum_{j=0}^{n_b} \begin{bmatrix} b_{j,n,1}(p(k)) & \dots & b_{j,n,n_u}(p(k)) \end{bmatrix} q^{-j}U(k) + e(k) \tag{5-7}$$

the scheduling dependent parameters are collected in:

$$\begin{bmatrix} \phi_{n,1} & \dots & \phi_{n,n_g} \end{bmatrix} \triangleq \begin{bmatrix} a_{1,n} & \dots & a_{n_a,n} & b_{0,n,1} & \dots & b_{n_b,n,1} & b_{0,n,2} & \dots & \dots & b_{n_b,n,n_u} \end{bmatrix} \tag{5-8}$$

with $n_g = n_a + n_u(n_b + 1)$, where each $\phi_{n,i}$ is a real function with static dependence on $p(k)$. Furthermore, each $\phi_{n,i}$ can be linearly parameterized as:

$$\phi_{n,i} = \theta_{n,i,0} + \sum_{j=1}^{s_i} \theta_{n,i,j}\psi_{n,i,j} \tag{5-9}$$

with $\{\theta\}_{n=1;i=1;j=1}^{n_y,n_g,s_i}$ being the unknown parameters to be estimated and $\{\psi\}_{n=1;i=1;j=1}^{n_y,n_g,s_i}$ are functions chosen by the user. For notation simplicity further on, we will denote $p(k)$ as p_k and we will remove the subscript n from all the parameters (i.e., ϕ_{n,n_g} will be denoted as ϕ_{n_g} and output $y_n(k)$ will be denoted as $y(k)$). After having each scheduling dependent coefficients linearly parameterized, lets introduce the unknown parameters θ denoted:

$$\theta = \begin{bmatrix} \theta_{1,0} & \dots & \theta_{1,s_i} & \theta_{2,0} & \dots & \dots & \theta_{n_g,s_{n_g}} \end{bmatrix} \tag{5-10}$$

and the regressor ψ is denoted by:

$$\begin{aligned}\varphi(k) \;=\; & \big[-y(k-1) \;\; -\psi_{1,1}(p_k)y(k-1) \;\; \dots \\ & -\psi_{1,s_1}(p_k)y(k-1) \;\; \dots \;\; -\psi_{n_a,s_{n_a}}(p_k)y(k-n_a) \\ & u_1(k) \;\; \dots \;\; -\psi_{(n_a+n_b+1),s_{n_b},1}(p_k)u_1(k-n_b) \;\; \dots \\ & u_{n_u}(k) \;\; \dots \;\; -\psi_{n_g,s_{n_b},n_u}(p_k)u_{n_u}(k-n_b)\big].\end{aligned} \tag{5-11}$$

The linear regression model of the MISO LPV-ARX structure is formulated as :

$$y(k) = \theta^\top \varphi(k) + e(k). \tag{5-12}$$

After the model have been written into the linear regression form, the unknown parameter vector θ can be estimated analytically from the identification data set $\mathcal{D}_N = \{(u(k), p(k), y(k))\}_{k=1}^N$. This is done by solving the least-squares problem:

$$\hat{\theta} = \arg\min_{\theta \in \mathbb{R}^n} \mathcal{V}(\theta, e) \tag{5-13}$$

under the cost function:

$$\mathcal{V}(\theta, e) \triangleq \frac{1}{N} \parallel e(k) \parallel_{\ell_2}^2, \;\; e(k) \triangleq y(k) - \theta^\top \varphi(k). \tag{5-14}$$

The cost function on Eq. 5-14 is also named as *quadratic criterion function* that leads to quadratic optimization problem, which generally has a unique solution that can be calculated analytically. In order to guarantee a unique solution of this least-squares problem, it is assumed that $\{\psi_{i,j}\}_{i=1,j=1}^{n_g,s_i}$ are chosen such that the system is globally identifiable (there exist no θ and θ'), such that the one step ahead predictor is not distinguishable for θ and θ', and that the data set provides a persistently exciting regressor (Eq. 5-12). This linear least-squares problem is solved by organizing the data set as:

$$Y = \big[y(1) \;\; y(2) \;\; \dots \;\; y(N)\big]^\top, \tag{5-15}$$

$$\Phi = \big[\varphi(1) \;\; \varphi(2) \;\; \dots \;\; \varphi(N)\big]^\top. \tag{5-16}$$

Then, the solution for estimated parameter $\hat{\theta}_N$ can be calculated by solving:

$$\hat{\theta}_N = (\Phi^\top \Phi)^{-1} \Phi^\top Y. \tag{5-17}$$

Although LPV system identification in the linear regression setting looks simple to be applied, there remains questions on the numbers of parameters θ_i and what kind of underlying functions $\psi_{i,j}$ are needed for a parsimonious model of the system. The choice of $\psi_{i,j}$ can range from simple polynomial basis to complicated rational or even discontinuous basis functions. Often a large set of basis functions $\{\psi_i, j\}$ are applied in the modeling, and from those large set, sometimes only a few are needed for an accurate approximation (Tóth et al. (2011)).

This means that most of the time the LS estimation will face an over-parameterized model estimation problem, where the underlying true parameter vector θ_0 can be sparse. This often results in a large variance of the parameters estimation. Moreover, if there is no prior information about the system nonlinearities, the set $\{\psi_{i,j}\}$ can be particularly inadequate, leading to potential structural bias. For the case study in this thesis, the above problem results in difficulties to apply global approaches of LPV system identification. Therefore, we propose to follow an alternative approach, which is to apply a non-parametric modeling approach by taking advantages of the attractive properties of *Support Vector Machines*'s (SVM) in particularly its least-square form.

5-2 LS-SVM

Support vector machines were originally developed as a class of supervised learning methods for efficient reconstruction of underlying functional relationships and structures in data (Vapnik 1995). The standard SVM takes a set of input data namely $x_i \in \mathbb{R}^d$ and label each one of them into two binary categories $y_i \in \{-1, +1\}$. The separation is done by constructing a linear classifier $\omega^\top x + b = 0$ corresponding to a hyperplane in the input space (denote ω as the normal vector of the hyperplane).

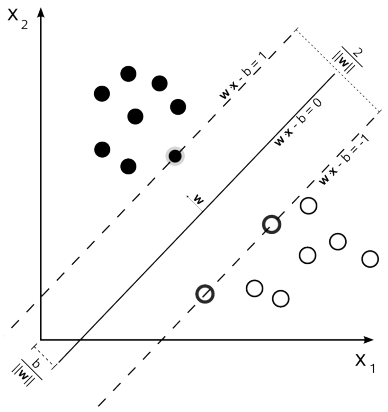


Figure 5-2: Illustration of SVM methodologies

An optimal classifier will separate the mapped data set by the highest margin ($\frac{2}{\|\omega\|_2}$). In many separation cases, the input data set is not always *linearly separable*. To do the separation, one have to construct a nonlinear classifier or/and introduce a relaxation of the constraints (*soft margin*). The relaxed constraints are applied in the margin optimization, if there exist no hyperplane that can split the input sets. These constraints will tolerate a misclassifications on the input data, by introducing slack variables ξ. On the other hand, for the nonlinear case the input data x_i will be mapped into higher dimension space (*feature space*) by introducing a feature map $\vartheta(.) : \mathbb{R}^d \to \mathbb{R}^{n_\mathcal{H}}$. The extension from linear SVM classifiers to nonlinear SVM classifiers is straight forward, but be aware that $\vartheta(x)$ can be infinite dimensional, and hence, also the vector ω. By considering these two conditions, the margin optimization of SVM's can be written as:

$$\begin{aligned} \min_{w,b,e} \quad & \mathcal{J}_p(\omega, e) = \frac{1}{2}\omega^\top\omega + c\sum_{k=1}^{N}\xi_k \\ \text{s.t.} \quad & y_k(\omega\vartheta(x_k) + b) \geq 1 - \xi_k, \quad k \in 1, \ldots, N \\ & \xi_k \geq 0, \quad k \in 1, \ldots, N \end{aligned} \tag{5-18}$$

where the objective function aims at achieving a trade-off between maximizing of the margin ($\frac{2}{\|\omega\|_2}$) and the amount of tolerated misclassification which are controlled by the regularization constant $c > 0$. The optimization written in Eq. 5-18 are usually called the *primal form* of the SVM. The solution in the primal form can become a hard problem if the dimensional space of the feature map became very high. Alternatively, the optimization problem of SVM can be solved in a different manner. Instead to explicitly declare the feature map $\vartheta(x_k)$, one can state the optimization problem in its *dual form* by using the *Karush-Kuhn-Tucker* (KKT) conditions for optimality. First, lets introduce the Lagrangian

$$\mathcal{L}(\omega, b, \xi; \alpha, \nu) = \mathcal{J}_P(\omega, \xi) - \sum_{k=1}^{N}\alpha_k(y_k[\omega^\top\vartheta(x_k) + b] - 1 + \xi_k) - \sum_{k=1}^{N}\nu_k\xi_k \tag{5-19}$$

with *Lagrangian multipliers* $\alpha_k \geq 0, \nu_k \geq 0$ for $k = 1, \ldots, N$. The optimal solution of the primal problem is obtained at the saddle point of the Lagrangian

$$\max_{\alpha,\nu}\min_{w,b,e}\mathcal{L}(\omega, b, \xi; \alpha, \nu). \tag{5-20}$$

Then, you can eliminate the primal variables by setting partial derivative of the Lagrangian with respect to the primal variables ω, b, and ξ_k to be equal to zero:

$$\begin{cases} \frac{\partial\mathcal{L}}{\partial\omega} = 0 \quad \rightarrow \quad \omega = \sum_{k=1}^{N}\alpha_k y_k\vartheta(x_i(k)), \\ \frac{\partial\mathcal{L}}{\partial b} = 0 \quad \rightarrow \quad \sum_{k=1}^{N}\alpha_k y_k = 0 \\ \frac{\partial\mathcal{L}}{\partial\xi_k} = 0 \quad \rightarrow \quad 0 \leq \alpha_k \leq c, \;\; k = 1, \ldots, N. \end{cases} \tag{5-21}$$

In the end, we arrive to the *dual problem* of the SVM:

$$\begin{aligned} \max_{\alpha} \quad & \mathcal{J}_D(\alpha) = -\tfrac{1}{2}\sum_{k,l=1}^{N} y_k y_l(\vartheta^\top(x_k)\vartheta(x_l))\alpha_k\alpha_l + \sum_{k=1}^{N}\alpha_k \\ \text{s.t} \quad & \sum_{k=1}^{N}\alpha_k y_k = 0 \\ & 0 \leq \alpha_k \leq c, \;\; k = 1, \ldots, N. \end{aligned} \tag{5-22}$$

This dual form is more attractive in the LPV identification point of view due to its capability to not have to solve the identification via classical overparametrization methods. This is because in this form, we can not only formulate the optimization problem without explicitly describing the normal vector ω, but also substitute the dot product of the feature map $(\vartheta^\top(x_k)\vartheta(x_l))$ by utilizing a trick known as the *Kernel Trick*.

Moreover, there are an extension of the SVM named Least-Square SVM(LS-SVM), where the original QP is simplified to an analytic solution of a linear regression problem and equality constraints (Suykens, 1999). This simplification is the one that will be adapted in to the LPV-ARX formulation. The original SVM is modified at two points. The first one, is that instead of using an inequality constrains, the LS-SVM takes equality constraints where the value of one at the right hand side is considered as a target value than a threshold value. For this target value, an error variable e_k will tolerate a misclassification in the case of overlapping distributions. These error variables play a similar role as the slack variables ξ_k in SVM formulations and are controlled by the regularization parameter γ. Second, a squared loss function is used for this error variable transforming the cost function into a so called *Ridge regression* form. These modifications will greatly simplify the problem.

The LS-SVM optimization problem is written as:

$$\begin{aligned} \min_{w,b,e} \quad & \mathcal{J}_p(\omega, e) = \frac{1}{2}\omega^\top\omega + \gamma\frac{1}{2}\sum_{k=1}^{N} e_k^2 \\ \text{s.t.} \quad & y_k[\omega^\top\vartheta(x_k) + b] = 1 - e_k, \quad k \in 1, \ldots, N \end{aligned} \tag{5-23}$$

This LS-SVM formulation can be solved in its dual form. Once more, the dual form can obtained by formulating the Lagrangian of (as seen in Eq. 5-19):

$$\mathcal{L}(\omega, b, e; \alpha) = \mathcal{J}_P(\omega, \xi) - \sum_{k=1}^{N} \alpha_k(y_k[\omega^\top\vartheta(x_k) + b] - 1 + e_k) \tag{5-24}$$

where the α_k values are the Lagrange multipliers, which can be positive or negative now due to the equality constraints. The KKT conditions for optimality are given as:

$$\begin{cases} \frac{\partial\mathcal{L}}{\partial\omega} = 0 \quad \rightarrow \quad \omega = \sum_{k=1}^{N} \alpha_k y_k \vartheta(x_i(k)), \\ \frac{\partial\mathcal{L}}{\partial b} = 0 \quad \rightarrow \quad \sum_{k=1}^{N} \alpha_k y_k = 0 \\ \frac{\partial\mathcal{L}}{\partial e_k} = 0 \quad \rightarrow \quad \alpha_k = \gamma e_k, \qquad\qquad k = 1, \ldots, N \\ \frac{\partial\mathcal{L}}{\partial\alpha_k} = 0 \quad \rightarrow \quad y_k[\omega^\top\vartheta(x_k) + b] - 1 + e_k, \quad k = 1, \ldots, N. \end{cases} \tag{5-25}$$

The value of α and b can be obtained by formulating a linear KKT problem:

$$\left[\begin{array}{c|c} 0 & y^\top \\ \hline y & \Omega + I/\gamma \end{array}\right] \left[\begin{array}{c} b \\ \hline \alpha \end{array}\right] = \left[\begin{array}{c} 0 \\ \hline 1_v \end{array}\right] \tag{5-26}$$

where $y = [y_1 \quad \dots \quad y_N]^\top$, $1_v = [1 \quad \dots \quad 1]^\top$, $e = [e_1 \quad \dots \quad e_N]^\top$, $\alpha = [\alpha_1 \quad \dots \quad \alpha_N]^\top$. The variable Ω is known as the *kernel matrix* which is constructed by

$$\begin{aligned} \Omega_{k,l} &= y_i(k)\vartheta_k^\top(x_k)\vartheta_l(x_l)y_i(l) \\ &= y_k K(x_k, x_l) y_l \qquad k,l = 1,\dots,N. \end{aligned} \tag{5-27}$$

Solving the linear KKT problem is the core problem in constrained nonlinear optimization problems in general. This linear KKT is solved efficiently for a given data set, however a particular feature of the solution is that the dimension of α grows with the number of training data N ($\alpha \in \mathbb{R}^N$) and it is not determined by the dimension of the input space.

The main advantage of LS-SVM is its computationally attractive properties. These approaches have also had a significant impact on nonlinear block model identification via various LS-SVM approaches (Suykens et al. 2002; Goethals et al. 2005; Falck et al. 2009; Giri and Bai 2010), and also on the LPV identification framework (Tóth et al. (2011)). In Tóth (2011), the LPV-LS SVM are formulated for the *Single-Input-Single -Output* SISO case and have been tested on a simple academic example. Within this thesis, we further extend the LPV LS-SVM formulation for MIMO case and application of this approach will be tested on the distillation column case study. Furthermore, due to the slow settling time and input restriction of the column, the LPV LS-SVM is also further enhanced to handle the large scale problem (See Section 5-3-2).

5-3 LPV modeling in the LS-SVM setting

First of all, in order to adopt the LS-SVM in the LPV identification setting, the structural dependence of model coefficients ϕ_i (Eq. 5-9) will be assumed to be a priori unknown. Then, the model equation (Eq. 5-5) is re-formulated in the similar manner as in the LS-SVM:

$$\mathcal{M}_{\omega,\vartheta} : y(k) = \sum_{i=1}^{n_g} \omega_i^\top \vartheta_i(p_k) x_i(k) + e(k), \tag{5-28}$$

where $\vartheta_i : \mathbb{R} \to \mathbb{R}^{n_H}$ denotes an undefined, potentially infinite dimensional feature map, and ω_i is the $i-$th parameter vector. Note that the model coefficient function is now described as $\phi_i(p_k) = \omega_i^\top \vartheta_i(p_k)$. The previous output, the current and previous inputs will be collected in the variables $\{x_i\}_{i=1}^{n_g}$, as:

$$x_i(k) = y(k-i), \quad \text{for} \quad i = 1,\dots,n_a, \tag{5-29}$$

$$x_{n_a+1+j}(k) = u_1(k-j), \quad for \quad j = 0,\dots,n_b, \tag{5-30}$$

$$x_{n_h}(k) = u_{n_u}(k-j), \quad for \quad j = 0,\dots,n_b, \tag{5-31}$$

with $n_h = n_a + n_u(n_b+1)$. The model equation written in Eq. 5-28, is then formulated in a regression form, by introducing:

$$\omega = \begin{bmatrix}\omega_1^\top & \dots & \omega_{n_g}^\top\end{bmatrix}^\top \in \mathbb{R}^{n_g n_H} \tag{5-32}$$

and the regressor:

$$\varphi(k) = \begin{bmatrix}\vartheta_1^\top(p_k)x_1(k) & \dots & \vartheta_{n_g}^\top(p_k)x_{n_g}(k)\end{bmatrix}^\top. \tag{5-33}$$

As a result, the linear regression form of the model equation becomes:

$$y(k) = \omega^\top \varphi(k) + e(k). \tag{5-34}$$

In this regression form, we can employ the technique available from the LS-SVM approach to achieve a non-parametric estimation of the model coefficients. The model coefficient identification is obtained via solving a similar least-square optimization problem as explained in the previous section. In the LPV LS-SVM settings, the optimization problem is written as:

$$\begin{aligned} \min_{w,e} \quad \mathcal{J}(\omega, e) &= \frac{1}{2}\sum_{i=1}^{n_g} \omega_i^\top \omega_i + \frac{\gamma}{2}\sum_{i=1}^{n_g} e^2(t_k) \\ \text{s.t.} \qquad e(k) &= y(k) - \sum_{i=1}^{n_g} \omega_i^\top \vartheta_i(p_k)x_i(k) \quad \forall k \in 1, \dots, N \end{aligned} \tag{5-35}$$

where the scalar $\gamma \in \mathbb{R}_0^+$ is the *regularization parameter*. This regularization parameter will allow some fitting error (identified model coefficient variance) in order to achieve an arbitrary small biased estimate of the model coefficient by increasing the regularization parameter value. It is also one of the parameters that need to be tuned in the identification routine. As we aim to solve this optimization problem in "non-parametric" approach, this so called primal form optimization problem needs to be formulated in its dual form. The dual form formulation would require us to construct the Lagrangian:

$$\mathcal{L}(\omega, e, \alpha) = \mathcal{J}(\omega, e) - \sum_{k=1}^{N} \alpha_k \Big(\sum_{i=1}^{n_g} \omega_i^\top \vartheta_i(k)x_i(k) + e(k) - y(k) \Big) \tag{5-36}$$

with $\alpha_k \in \mathbb{R}$ being the Lagrangian multipliers. The Lagrangian have solutions which are located in its saddle point. This point is obtained by minimizing the Lagrangian with respect to ω and e, but maximize it with respect to α (See Eq. 5-20). In other word, it can be calculated by having a partial derivative of the Lagrangian with respect to the prediction error e, parameter vector ω and lagrangian multiplier α to be equal to zero:

$$\begin{cases} \frac{\partial \mathcal{L}}{\partial e} = 0 \quad \rightarrow \quad \alpha_k = \gamma e(k), \\ \frac{\partial \mathcal{L}}{\partial \omega_i} = 0 \quad \rightarrow \quad \omega_i = \sum_{k=1}^{N} \alpha_k \vartheta_i(k)x_i(k), \\ \frac{\partial \mathcal{L}}{\partial \alpha_k} = 0 \quad \rightarrow \quad e(k) = y(k) - \sum_{i=1}^{n_g} \omega_i^\top \vartheta_i(k)x_i(k). \end{cases} \tag{5-37}$$

The substitution of the first two equation in Eq. 5-37 into the third part, lead to *dual form* of Eq. 5-35 in terms of solving:

$$y(k) = \sum_{i=1}^{n_g} \underbrace{\left(\sum_{k=1}^{N} \alpha_k x_i(k) \vartheta_i^\top(k) \right)}_{\omega_i^\top} \vartheta_i(k) x_i(k) + \underbrace{\gamma^{-1}\alpha_k}_{e(k)} \tag{5-38}$$

for $k \in \{1, \ldots, N\}$. As explained in the SVM section, this dual form is the "non-parametric" counterpart of the primal form. This means that neither the parameter vector ω or the feature map ϑ_i are need to be computed directly. Instead, we utilize a trick known as the *Kernel trick*, where the dot product $\vartheta_i^\top(j)\vartheta(k)$ can be substituted with a kernel $K^i(p(j), (p(k)))$. This substitutions is highly desirable, because we are aiming at solving the LPV identification problem from a function estimation perspective rather than doing the traditional over-parametrization approach for the LPV identification using priori selected basis.

5-3-1 Kernel trick

The principle idea behind the kernel trick is to introduce a kernel function $K^i(p(j), (p(k)))$ which give a similarity measures between $\vartheta_i(k)$ at time k with respect to all other $\vartheta_i(j)$ in the data set. Furthermore, this measure replaces the cross products of the unknown (undefined) feature map. The kernel trick enables us to work in huge dimensional feature spaces without actually having to do explicit computations in this space. However, in order to obtain a unique solution for the dual form, the kernel is defined as a positive definite function. By employing the kernel, the Lagrangian multipliers α in the dual form (Eq. 5-38), can be calculated by re-writing it into:

$$Y = (\Omega + \gamma^{-1} I_N)\alpha \tag{5-39}$$

where $\alpha = \begin{bmatrix} \alpha_1 & \ldots & \alpha_N \end{bmatrix} \in \mathbb{R}^N$, and the *Kernel matrix* Ω is defined by:

$$[\Omega]_{j,k} = \sum_{i=1}^{n_g} [\Omega^i]_{j,k} \tag{5-40}$$

with

$$\begin{aligned} [\Omega^i]_{j,k} &= x_i(j)\vartheta_i^\top(j)\vartheta_i(k)x_i(k) \\ &= x_i(j)\langle \vartheta_i(j)\vartheta_i(k)\rangle x_i(k) \\ &= x_i(j)(K^i(p(j), p(k)))x_i(k). \end{aligned} \tag{5-41}$$

Note that the superscripts on both the kernels K^i and the kernel matrices Ω^i are used only for numbering purpose for the related model coefficient ϕ_i. Consequently, the set of K^i's defines Ω and characterizes the feature maps $\{\vartheta_i\}_{i=1}^{n_g}$ in an efficient fashion. This allows to characterize a wide range of nonlinear dependencies as a linear combination of infinitely many functions defined through the choice of the particular inner product and a relatively low dimensional

parameter α. There are many kernel types such as: linear, polynomial, Radial Basis Function (RBF), multiquadratic, etc. The choice of a Kernel depends on what we are trying to model. A polynomial kernel, for example, allows us to model feature conjunctions up to the order of the polynomial. Radial basis functions allow to pick out circles (or hyperspheres) in contrast with the Linear kernels, which allow only to pick out lines (or hyperplanes). Choice of the most appropriate kernel highly depends on the problem at hand and fine tuning its parameters can easily become a tedious and cumbersome task. Searching for different kernels either via trial and error or other exhaustive means can be a computationally daunting problem (Chapelle & Vapnik, 1999). Therefore, we decide to focus on the most used kernel in nonlinear function estimation, which is the *Radial Basis Function* (RBF) kernel:

$$K^i(p_j, p_k) = \exp\left(- \frac{\| p_j - p_k \|_2^2}{\sigma_i^2} \right) \tag{5-42}$$

with σ_i specifies the width of the RBF kernels, and p_j and p_k are the scheduling variables at time j and k. Along with the regularization parameter γ, the kernel width σ is one of the tuning parameter that will determine the trade off between bias-variance of the identified model (See Section5-4-1). These two parameters are usually called *hyper parameters*.

With the kernel trick at hand, we will be able to solve the optimization problem written in Eq. 5-38 in a non-parametric manner. The first step is to estimate the model coefficient of the LPV model estimate is to compute the Lagrangian multipliers α from the training data set:

$$\alpha = (\Omega + \gamma^{-1} I_N)^{-1} Y. \tag{5-43}$$

Afterwards, the Lagrangian multiplier can be used to estimate the model coefficients as:

$$\phi_i(.) = \omega_i^\top \vartheta_i(k) = \sum_{j=1}^{N} \alpha_j x_i(j) K^i(p(j), .). \tag{5-44}$$

Since this is a non-parametric modeling approach, each value of the scheduling variable $p(k)$ of the data set $\mathcal{D}_N = \{(u(k), p(k), y(k))\}_{k=1}^N$ (which is used for the model coefficient identification purpose) will provide local information to estimate the function $\phi_i(.)$ as the linear combinations with lagrangian multipliers α_i .In order to obtain a model coefficient value at any desired scheduling variable value, we have to do the Eq. 5-44 at each scheduling variable.

This identification approach will only produce unbiased estimates if the underlaying system really have an ARX noise structure. Even if the system does not have the ARX noise structure, by choosing a large regularization parameter γ we can make the regularization bias to be more dominant than the noise bias. Of course there are more we can to overcame this more advanced approach such as IV based SVM that can handle this bias (Laurain et al. 2011).

5-3-2 Large Scale Problem

The dimension of the Lagrangian multiplier α_i and its corresponding kernels K^i, depend on the size of the data set $\mathcal{D}_N = \{(u(k), p(k), y(k))\}_{k=1}^N$. The size of the lagrangian multiplier will

be equal to the length of data set N, and the kernel matrix will be as huge as $N \times N$. For a fast low order system, small data set might already be enough to adequately estimate the model coefficient. However, a slow large scale system such as the high purity distillation column will require a large data set to achieve an accurate estimate of the coefficient functions. Hence, it is computationally inefficient to estimate the model coefficients without any modification of the previously described procedure. When facing this problem, we can separate the data set into numbers of data batches (n_{batch}) with a smaller size. The final estimate of the model coefficient will be the average of the estimated model coefficients ϕ_i^* that are calculated in each smaller data batches.

$$\phi_i^* = \frac{1}{n_{\text{batch}}} \sum_{n=1}^{n_{\text{batch}}} \phi_i, n. \tag{5-45}$$

When deciding to use this modification, we have to make sure that the dynamics of the system are well captured in each data batches. If the input excitation used in the data set is a white noise excitation, this will not give inaccurate estimation on the average of the model coefficients. However, we have to remember that the distillation column have a limitation on the input excitation that can be used for the system. Namely it is not allowed to change rapidly between operating conditions as the vapor and liquid flows in the column will start to significantly vary and possibly collapse, resulting in a turbulent unwanted behavior. This limitation will lead to further modification of the procedure to choose each member of the data set at a randomly chosen time from the original data set $\mathcal{D}_N$. This way the new batch of the data set will surely capture most of the system dynamics that have been collected in the original data set. As the calculation of the model coefficients also depend on past inputs and outputs value, for each randomly chosen time τ, those past values have to be captured as well. The previous input and output values that need to be captured, will depend on the model coefficient functions n_a and n_b. The past values of scheduling variable are not needed since the scheduling dependencies of the system are static (the system only depend on the instantaneous value of the scheduling variable). With this modification, we can reduce the computational load of the SVM approach and ensure the accuracy of the estimated model coefficients.

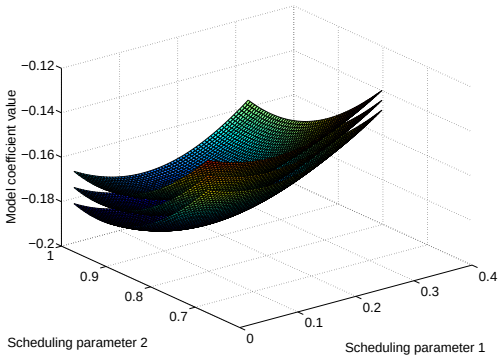
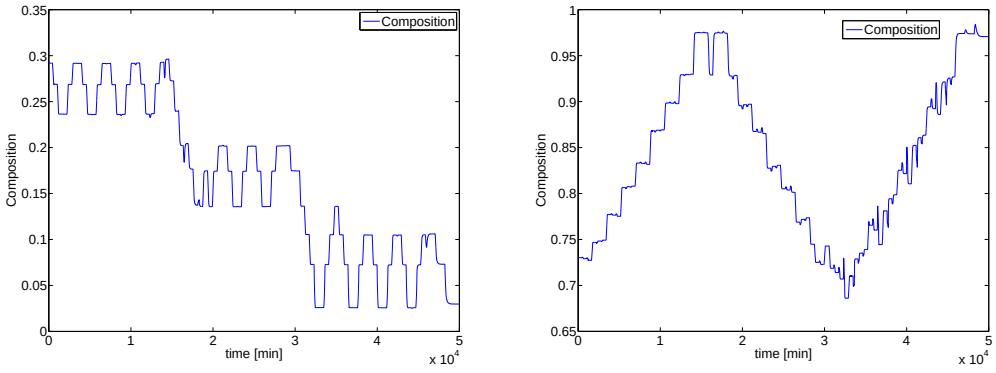


Figure 5-3: Effect of LPV LS-SVM modification for large scale problem on the estimated model coefficient (and its standard deviation).

5-4 Result and Analysis

Based on the theories explained in this chapter, the SVM based global LPV identification approach is then applied in the distillation column case study. In order to model a system as an LPV model, we have to select the scheduling variables that will be used for the model. As explained in Section 3-1-1, the scheduling variables can be any information coming from the system as long as they are able to uniquely describe the changing dynamics of the system itself. The same scheduling variable that are used in the local LPV identification approach (bottom product composition and top product composition), can still be used for the identification in the LPV LS-SVM setting. However, we would like to demonstrate the flexibility of the scheduling variable selection, hence, we will select a different scheduling just to show the flexibility of scheduling variables selection. The scheduling variables that we choose for the LPV LS-SVM identification, came from the composition of the light component on the 10th and 100th tray level. 5-4



(a) First scheduling variable **(b)** Second scheduling variable

Figure 5-4: Scheduling variable for global identification approaches

In reality, we do not have a lot of options to choose the scheduling variable, due to the limited availability of the system signals that are captured with the installed sensor. However, since we are using the high fidelity model of the distillation column, we can use the freedom to take any varying parameters that comes from the distillation column nonlinear model. These scheduling variables, along with the corresponding system inputs and outputs will be gathered together in the data set used identification ($\mathcal{D}_N$) (with 50000 data point), identification validation (used to tune the hyper parameters)($\mathcal{D}_O$), and an independent dataset used for validation($\mathcal{D}_V$). The data set can be seen in appendix B

In Section 2-3-1 we have decided that we will only consider measurement noise for this system. To further demonstrate the effectiveness of the LPV LS-SVM methodology, three different noise conditions such as low noise condition with 40dB SNR, medium noise condition of 25dB SNR, and severe noise condition of 10dB SNR (See Equation 4-23) of white noise will be added in the output signal. The variance of the noise is scaled with respect to magnitude of the distillation column output to produce the specific SNR mentioned above.

From the noise setting, it is seen that the identification using LPV-ARX model structure will result in a bias estimates. We can overcame this bias with more advanced approach such as

IV based SVM (Laurain et al. 2011), however, the method is yet to be implemented for this project. Additionally, the identification result shown further in this section show that good model estimates of the system can still be achieved via tuning of the hyper parameter.

The experiment on LPV model identification in LPV LS-SVM setting is done in a monte carlo study using 100 attempt with only the measurement noise in the output of the system which are independently generated for the identification data set.

5-4-1 LPV LS-SVM identification procedure

Before conducting the identification experiment, lets summarize the steps of the identification cycle with respect to LPV LS-SVM based identification:

1. Collect identification data set($\mathcal{D}_N$), validation data set for hyper parameter tuning purpose ($\mathcal{D}_O$), and an independent validation data set($\mathcal{D}_V$).
2. Decide the model order (n_a and n_b)
3. Optimize the hyper parameters with respect to the LPV model output result based on the validation data set for the hyper parameter tuning purpose ($\mathcal{D}_O$).

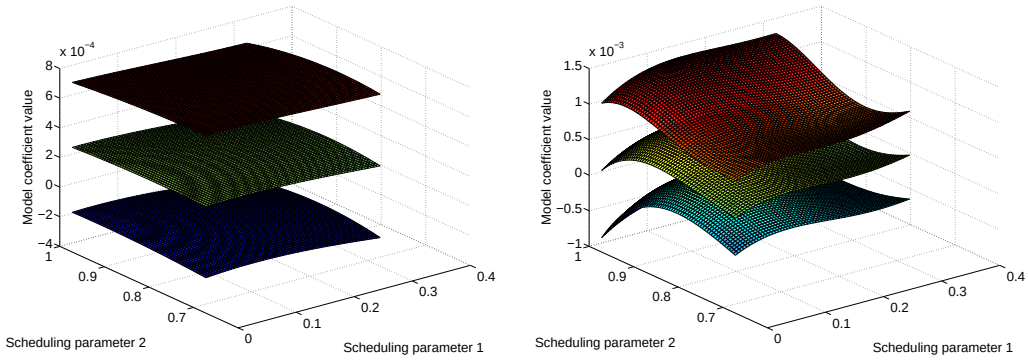
The LPV LS-SVM obtain the estimates of the model coefficients via solving a least-squares optimization problem. Therefore, the first step on conducting the identification is to obtain a rich identification data set which contain the dynamics of the distillation column. For this purpose, the input excitation signal should be designed such that it travel through most part of the system operating regime (See Section 2-3-1). The input and output signal used for the identification can be seen in Figure 2-7, this input and output signal, correspond to the scheduling the corresponding scheduling variable can be seen in Figure 5-4.

On this experiment, because we are actively tuning the hyper parameters to achieve an accurate model of the system, we require a validation data set for the identification purpose ($\mathcal{D}_O$) , and another independent data set to validate LPV model obtained from the experiment ($\mathcal{D}_V$). The independent data set is necessary to make sure that the hyper parameter tuning is not tuned to achieve the best fit for this validation data set.

The second step of the procedure, is to determine the order of the LPV model (n_b and n_a). Based on the previously gathered information in the local LPV identification approach, we know that a fourth order LTI model is enough to describe the local dynamics of distillation column in the vicinity of an operating point. This information leads to the initial guess of the LPV system order choices of the output and input side ($n_a = 4$ and $n_b = 4$). Although further in this section we will also investigate different model orders.

The next items that will influence the quality of the identified LPV model are the *hyper parameters* of the LS-SVM which are the regularization parameter (γ) and width of the RBF kernel (σ). The procedural steps from the third step until the sixth step are dedicated to obtain the optimal hyper parameters that will lead to the best training validation result. The value of the regularization parameter will give a trade-off between bias and variance of the model, where small value of regularization parameter will give larger bias but more smooth estimates. For the width of the RBF kernel (σ), this value will represent range for the

similarity measure of the kernel. As seen on Figure 5-5, the hyper parameter tuning will also influence the shape of the estimated coefficient functions surface (coefficient function value in the operating regime of the system). By increasing the kernel width σ, the surfaces of the estimated functions flatten out due to the fact that it models less noise. The flattened surface will announce more local estimation errors. The increase of regularization parameters have the opposite effect, it will give more fluctuation because it will aim to fit the model better.



(a) High σ **(b)** Low σ

Figure 5-5: Effect of hyper parameters tuning on the estimated model coefficient (and its standard deviation)

The tuning of these hyper parameter can be tricky and time consuming. One can manually do a cross sectional tuning such as bi-section search to obtain the optimal combination of the hyper parameters. Since the noise in the setup is not in the ARX setting, we would already expect bias caused by the noise. Therefore, bigger regularization parameters value ($\gamma \simeq 10^5$) are required to lessen this bias. The search of the optimal parameters can be narrowed in the region that have better fit ratio. The main disadvantage of this cross sectional tuning procedure is that the search might end in a local optimum if the tuning surface is not linear. Other things to keep in mind about the tuning of the hyper parameters, is that with different noise conditions affecting the distillation column output, it is necessary to conduct the tuning procedure in each of those noise conditions.

5-4-2 Result of the LPV LS-SVM identification for the high-purity distillation column

After completing the identification procedure, the simulation error of the identified LPV model with respect to different noise conditions can be seen in the Table 5-1 and 5-2:

The increase of the variance of the noise influences the obtained accuracy of the model. However, it can be seen that the global LPV identification approach using the LPV LS-SVM algorithm have a decent result even under severe noise conditions. As the noise grows, the tuning of the hyper parameters will change. The width of the kernel σ will have an optimal value in bigger σ because we have to model less the noise and the flattened surface will announce more local estimation errors. The effect of this tuning will be seen as a spike in the

Table 5-1: Comparison of simulated LPV model 1st output (bottom product composition) fit ratio under different noise condition

	BFR	MSE	STD
Noiseless Training	99.96%	$2.34.10^{-10}$	$3.02.10^{-6}$
Noiseless Validation	99.87%	$3.38.10^{-9}$	$9.97.10^{-6}$
Low Noise (40dB) Training	99.74%	$1.12.10^{-8}$	$9.09.10^{-5}$
Low Noise (40dB) Validation	99.55%	$4.75.10^{-8}$	$1.54.10^{-4}$
Medium Noise (25dB) Training	99.49%	$4.04.10^{-8}$	$1.89.10^{-4}$
Medium Noise (25dB) Validation	99.07%	$2.09.10^{-7}$	$3.24.10^{-4}$
Heavy Noise (10dB) Training	98.41%	$3.93.10^{-7}$	$3.80.10^{-4}$
Heavy Noise (10dB) Validation	96.89%	$2.42.10^{-6}$	$6.35.10^{-4}$

Table 5-2: Comparison of simulated LPV model 2nd output (top product composition) fit ratio under different noise condition

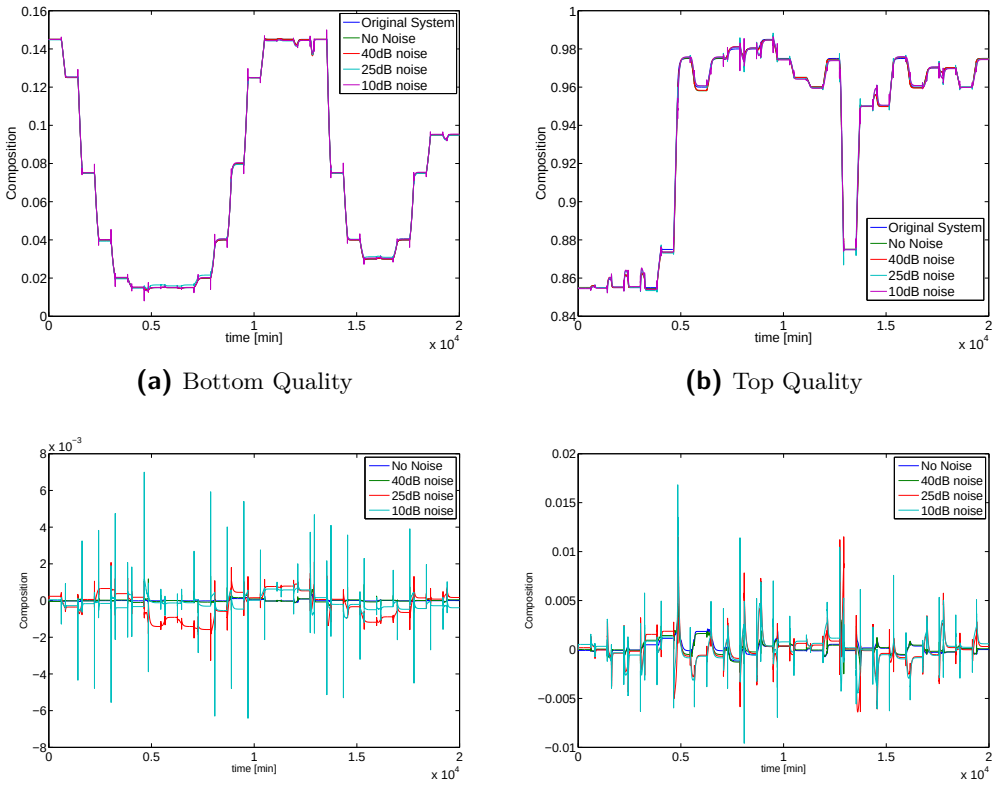
	BFR	MSE	STD
Noiseless Training	98.88%	$1.2813.10^{-7}$	$4.52.10^{-5}$
Noiseless Validation	98.35%	$6.35.10^{-7}$	$1.49.10^{-4}$
Low Noise (40dB) Training	98.71%	$1.71.10^{-7}$	$1.15.10^{-4}$
Low Noise (40dB) Validation	98.18%	$7.74.10^{-7}$	$2.43.10^{-4}$
Medium Noise (25dB) Training	97.79%	$5.02.10^{-7}$	$2.69.10^{-4}$
Medium Noise (25dB) Validation	96.98%	$2.14.10^{-6}$	$5.11.10^{-4}$
Heavy Noise (10dB) Training	95.78%	$1.80.10^{-6}$	$5.04.10^{-4}$
Heavy Noise (10dB) Validation	94.25%	$7.85.10^{-6}$	$8.16.10^{-4}$

transient dynamics of the LPV model output response. Hence, this response is not seen in when we choose small value of σ in low noise conditions.

Next we investigate the effect of choosing different model order for the identification of the LPV model. This is to make sure that we already make a right choice and also to try wether it is necessary to reduce or increase the model order. It would be very desirable if we can model the system with lower complexity by choosing lower model order, or in the other case, it might be necessary to increase the complexity in order to gain significant increase on the identification quality. The result shown in the Table 5-3 and Table 5-4 are the optimal result found after hyper parameters tuning in the noiseless environment.

Table 5-3: Comparison of simulated LPV model 1st output (bottom product composition) fit ratio under different model order

	BFR	MSE	STD
$n_a = 2; n_b = 2$ Training	99.92%	$7.54.10^{-10}$	$4.00.10^{-6}$
$n_a = 2; n_b = 2$ Validation	99.86%	$4.54.10^{-9}$	$1.05.10^{-5}$
$n_a = 6; n_b = 6$ Training	99.93%	$7.04.10^{-10}$	$3.66.10^{-6}$
$n_a = 6; n_b = 6$ Validation	99.88%	$3.19.10^{-9}$	$1.11.10^{-5}$



(a) Bottom Quality

(b) Top Quality

(c) Error on Bottom Product Composition

(d) Error on Top Product Composition

Figure 5-6: Comparison of the true and the simulated LPV model output under different noise condition

Table 5-4: Comparison of simulated LPV model 2nd output (top product composition) fit ratio under different model order

	BFR	MSE	STD
$n_a = 2; n_b = 2$ Training	98.44%	$2.44.10^{-7}$	$5.16.10^{-5}$
$n_a = 2; n_b = 2$ Validation	98.20%	$7.57.10^{-7}$	$1.90.10^{-4}$
$n_a = 6; n_b = 6$ Training	98.23%	$3.15.10^{-7}$	$5.63.10^{-5}$
$n_a = 6; n_b = 6$ Validation	98.26%	$7.08.10^{-7}$	$1.93.10^{-4}$

This test shows us that the output of the LPV model with second order model and sixth order model have slightly worse fit if compared to the result with the fourth order LPV model shown in Table 5-1 and 5-2. For the higher model order the worse result can be seen because we use force to use more model order than it was required which results in added dynamics that was unnecessary. For the lower model order result however, with this insignificance difference on the result, it might be necessary to reduce the model order to reduce the complexity and the computational effort of the LPV LS-SVM algorithm. However, the effect of changing the model order in different noise condition are yet to be further explored.

5-4-3 Comparison with overparametrization method

As the LPV LS-SVM provide an alternative model coefficient estimation over the classical over-parametrization method, it would be wise to also conduct comparison with the over-parametrization setting. The LPV-ARX model used in this method are constructed with the choice of polynomial basis $\psi_{i,j}$ and use the linear regression methodology explained earlier in this chapter (From Eq. 5-9 to Eq 5-17). The simulation result can be seen in Table 5-5 and 5-6.

Table 5-5: Comparison of simulated LPV model (identified via overparametrization method) 1st output (bottom product composition) fit ratio under different noise condition.

	BFR	MSE	STD
Low Noise (40dB) Training	99.87%	$3.48.10^{-6}$	$2.63.10^{-5}$
Low Noise (40dB) Validation	99.80%	$9.26.10^{-9}$	$3.64.10^{-5}$
Medium Noise (25dB) Training	99.72%	$1.59.10^{-8}$	$8.11.10^{-5}$
Medium Noise (25dB) Validation	99.59%	$3.93.10^{-8}$	$1.08.10^{-4}$
severe noise (10dB) Training	99.20%	$1.31.10^{-7}$	$2.52.10^{-4}$
severe noise (40dB) Validation	98.83%	$3.28.10^{-7}$	$3.46.10^{-4}$

Table 5-6: Comparison of simulated LPV model (identified via overparametrization method) 2nd output (top product composition) fit ratio under different noise condition.

	BFR	MSE	STD
Low Noise (40dB) Training	97.83%	$1.01.10^{-9}$	$2.67.10^{-5}$
Low Noise (40dB) Validation	98.43%	$5.73.10^{-6}$	$3.62.10^{-5}$
Medium Noise (25dB) Training	97.76%	$1.07.10^{-6}$	$7.82.10^{-5}$
Medium Noise (25dB) Validation	98.32%	$6.55.10^{-7}$	$1.09.10^{-4}$
severe noise (10dB) Training	97.67%	$1.16.10^{-6}$	$2.42.10^{-4}$
severe noise (40dB) Validation	97.89%	$1.04.10^{-6}$	$3.19.10^{-4}$

From Table 5-5 and 5-6, we see that the simulation result show a good fit with respect to the original system. This is because the correct basis choice which is polynomial. This basis choice might be true for the case of high fidelity model that we build. However, in real distillation column system the nonlinearity in the system might be more than the assumption that we include in the modeling of the high-fidelity model. By using LPV LS-SVM the prior nonlinearities information of the system is not necessary, therefore we do not need to worry about the basis selection.

5-5 Summary on the global LPV identification approach

When we are identifying a MIMO system such as distillation column, we can always decouple the system and identify it as a multiple multi input single output system, especially if we want to simplify the estimation problem. This treatment however, have to be handled carefully

and the model coefficient identification of the separated MISO system have to be conducted with the same data sets.

The classical ways to estimate the LPV-ARX model coefficients was to select a large set of basis function to describe the dynamics of the identified system. The classical overparametrization method might result in higher computational load and will result in a large variance of the parameters estimation. This further result in poor identification quality of the LPV system. Alternatively, in this thesis study, we are doing a non-parametric modeling approach, where we take advantages of supervised machine learning method to analyze data and recognize patterns. Upon doing the identification in the LPV LS-SVM framework, we do not need to address the nonlinearity of the system before estimating the model coefficients as the algorithm itself that will identify the system nonlinearities. This thing can be quite helpful if we did not have a prior information about system non-linearities. Moreover, the result shown from the experiments have shown the visibility of implementation of this algorithm in the distillation case study. The LPV LS-SVM methodology have shown that it can still works extremely well even under severe noise conditions.

Chapter 6

Conclusion and Recommendations

In the second and the third chapters of this report, Distillation column operation as well as LPV system framework have been discussed. It was also pointed out and motivated that an Linear Parameter Varying identification framework could be the efficient modeling solution for the process system in general, and high purity distillation column in specific. With this in mind, the fourth and the fifth chapters of the report were devoted on the development of separate approaches that try to capture the column dynamic from different perspectives of LPV identification approach. Hence the conclusions of each chapter were specific to the chapters only.

This chapter gives the overall conclusions and our recommendations on this thesis project. It summarizes the work and results obtained mainly in the previous two chapters. And afterwards some of our recommendations for further work are given to close this chapter.

6-1 Conclusions

The main goal of this thesis is to found efficient modeling solutions for high-purity distillation columns. For a system with high complexity such as distillation column, a first principle based modeling is often undesired due to the tendency of the modeling procedure to became very laborious and expensive. Data-driven identification based model are often became a solution instead of the first-principle based modeling since it only require the dynamics of the system to be captured in the data set used for identification. Hence, it provides reliable result which capture the dynamics of the system in its latest physical properties. However, the commonly practiced data-driven identification based modeling, i.e., PEM identification, have difficulties if the system is known to have nonlinear behavior in its operating regime. The identified linear model of the system will give valid response in the close vicinity of the system operating point, hence it cannot covers the whole operating regime. With these problem at hand, we propose LPV identification which offers low complexity and low cost modeling techniques as the solution. In order to shows the capabilities of LPV identification approach to model process system such as high purity distillation column, we have done several things such as:

1. Build a high-fidelity distillation column model that can represents the behavior of a high purity distillation column.
2. Experiment various LPV identification approaches on both local and global perspective identification for the distillation column system
3. Obtain results which convince us that the LPV can become an efficient modeling solution for the distillation column.

Upon conducting the LPV identification approach on distillation column setup, we have found several problem which solutions are not found in the available literatures. Therefore, we have made modification on the specific LPV identification approaches, such that the methodology became applicable for the high-purity distillation column case. The modification includes the separation of RBF parameter optimization on the local LPV identification case, dividing the estimation problem of LPV LS-SVM methodology in to smaller problem, and other else. More on this will be described further in the specific conclusion on each local and global LPV identification approach.

Things to note is that we have to decide which scheduling variable that will be used for the LPV modeling approach that we use. After we have chosen a specific variables for scheduling, we have to consistently use it for the whole identification approach, as the choice of the scheduling variables define the possible dynamical behaviors of the produced LPV model structure, hence, the representation capability of the model. This scheduling variable should be able to display the changing dynamics of the system or known as the changing operating point of the system. As said in 2-3-1 the operating points of the distillation column are usually denoted by the composition of the output (Top or bottom product). This composition can be measured via composition sensor that are placed in both product flows. Other parameters which vary as the operating points of the system changes, such as the composition or temperature measured on other trays can also be used as the scheduling signal. However, the choice of the scheduling variable should also consider applicability on the column itself. Through this thesis we have show the flexibility to chose the scheduling variable with different scheduling variables are chosen for the local and global identification approach.

In the end, the results and the experiments shown in this thesis, have shown the applicability of different LPV modeling and identification approaches for the high purity distillation case study.

6-1-1 Local LPV identification approach

The local LPV identification have a flexibility for not having to do an experiment which excite the system dynamics over its operating regime (Figure 2-7). This is beneficial if we do not have full access / permission to conduct such a test. By knowing the exact point where the system will be operated, we can efficiently model the system around those point, and interpolates them to give broader description of system dynamics in the operating regime.

Additionally, for the interpolation based on input/output schemes, if we are able to excite the system such that it give the dynamical behavior around the operating regimes, we could further optimize the quality of the derived LPV model. Several interpolation methods such

as polynomial and RBF based interpolation are able to do this optimization. The RBF based interpolation method have a flexibility to chose wether we want to do optimize the parameter of the interpolation weights or not. If we decide to do the optimization on the weight parameters, we could separate the multi dimensional RBF weight into multiple single dimension function and optimize them separately. This is done to lower the complexity of the optimization and to avoid some areas on the operating regimes to be poorly / overly weighted. On the polynomial interpolation, the optimization data set $\mathcal{D}_O$ are obligatory in order to obtain the overall polynomial weight. This will result in a precise weight of each LTI model. However, since the weight optimization is done to fit an output response of the model, the frequency response of the model will have a bad description over the operating regime of the system. This will give difficulties for the control design of the distillation column. Much straightforward interpolation method such as bilinear interpolation, does not have ability to do this. In return, the interpolation are much easier to be done in this manner.

Moreover, the behavior of the LPV model is highly dependent on the quality of the identified LTI models. This can be seen by the huge differences in the simulated output result when the interpolation are applied using LTI model obtained via linearization instead the PEM identification. From the preliminary experiments using LTI model obtained via linearization of the system we found that coefficient interpolation with polynomial based interpolation method have the best result in terms of output response and frequency response. It is shown that the resulting But it is also shown that the coefficient interpolation are more susceptible to stability issue when the identified model have large relative error. In the other hand the output interpolation scheme are less susceptible to this problem as long as the identified model are stable. However, the sensitive and oscillation behavior of the local model will still be inherited to the derived LPV model.

6-1-2 Global LPV identification approach

On the other hand, if we have the access to conduct experiment to excite the dynamical response of the system, the global LPV identification approach is more likely to be conducted if compared to the local LPV identification approach. In this manner we do not have to conduct identification on each desired operating point of the system. Instead, the identification can be solved in single procedure (with the exception of model re evaluation for the tuning of hyper parameter).

In this thesis, we have shown that alternative method in 'non parametric' manner can give good estimation result without any prior information of system nonlinearities. Moreover, We have made modification from the current method of LPV LS-SVM, such that the identification are able to handle MIMO system such as distillation column. We also propose a way to handle large-scale system by reducing the optimization problem in to smaller batch data set which each member are randomly chosen from the original identification data set. We have shown that the identified model coefficient by using this modification have small standard deviation on single identification run, and the monte carlo run still show small standard deviation for 100 identification attempt. We have also compared the result with the classical overparametrization method. This parametric manner is not to "overparametrized" since we only put polynomial basis. Hence, the simulation result of the LPV model still show a good result with respect to the validation data set. However, this also show that the non

parametrical manner can even obtain competing result even without prior setup for the system nonlinearities.

6-2 Recommendations

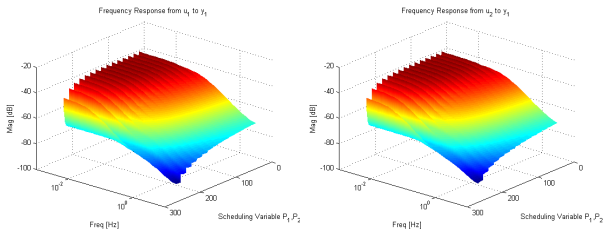
During this thesis works we have some recommendations on the future works that can be explored to better show the capability of LPV identification on the distillation column.

- Find better ways to identify local dynamics of the system.
  As we have difficulties on the PEM identification procedure, we might extend the identification into frequency domain identification of the local system behavior. This way we might get better result from the output response of the simulated LPV model.
- To use more realistic noise condition in the study.
  Although this need a specialist knowledge of the high purity distillation column operation, more realistic noise could better show the applicability of LPV in the real world scenarios. More over, by using different noise on the system this would induce bias to the current applied identification technique. However, we can apply more advanced methodology by using *Instrumental Variable* based identification in both local and global based identification.
- Manually assign the operating point of the local model used for interpolation.
  Although we have seen that the local identification approach by using the linearized model works well with gridding of 15x15 in the operating regime, we can further optimize (reduce) the number of the model by specifically place more model in the point where the system dynamics is more nonlinear (e.g in the high purity region).
- Improve the LPV LS-SVM identification procedure.
  This procedure can be further improved putting algorithm to select the kernel of the SVM and also to tune the hyper parameter automatically in the loop of the identification. Although this would also mean that we would increase the complexity of the optimization problem.
- Compare the LPV LS SVM with similar nonparametric approaches.
  We can also compare the LS-SVM that are chosen w.r.t to similar non-parametrical method such as disperse NNG, etc, this can provide more diversity off the option, but can also see if there are also problem on the other side. and we can also compare the computational load.
- Extend the study into control synthesis for the model.
  The promising result of LPV identification can be further extended by designing a controller for the high fidelity model. This way we can also show the usefulness of LPV model identified in the study.

Appendix A

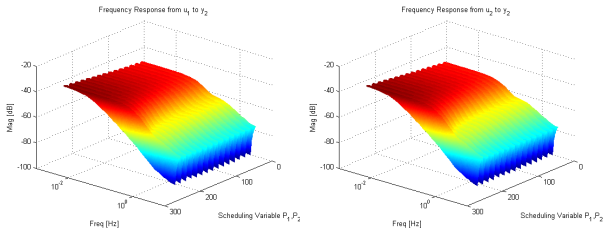
Frequency response figures in local LPV identification approach

These are figures about frequency response of the LPV model obtained from the preliminary experiment from local identification approach (Section 4-3-1). The Y axis show a combination of scheduling variable which are stacked together. The order of scheduling variable are chosen such that it translates from low purity region earlier on Y axis, towards high purity region at the higher value of that axis.



(a) From 1st input to 1st output

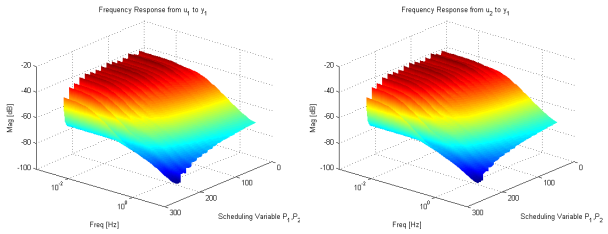
(b) From 2nd input to 1st output



(c) From 1st input to 2nd output

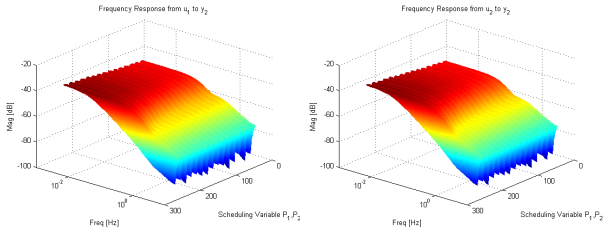
(d) From 2nd input to 2nd output

Figure A-1: Frequency response of LPV model with coefficient interpolation using RBF interpolation method



(a) From 1st input to 1st output

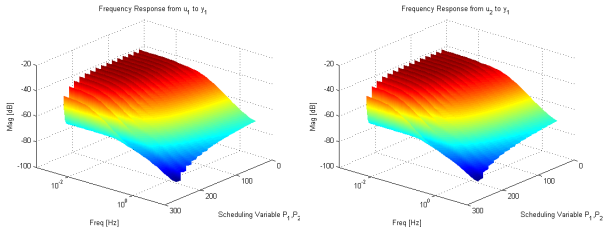
(b) From 2nd input to 1st output



(c) From 1st input to 2nd output

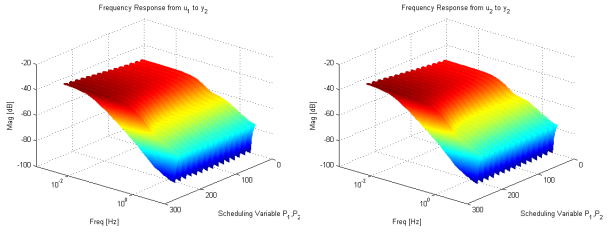
(d) From 2nd input to 2nd output

Figure A-2: Frequency response of LPV model with coefficient interpolation using polynomial interpolation method



(a) From 1st input to 1st output

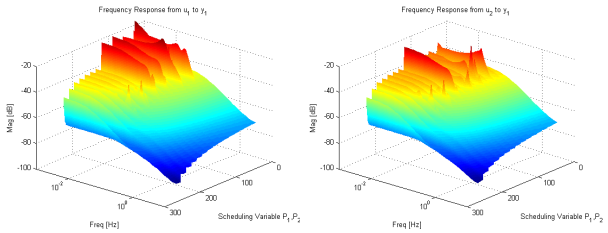
(b) From 2nd input to 1st output



(c) From 1st input to 2nd output

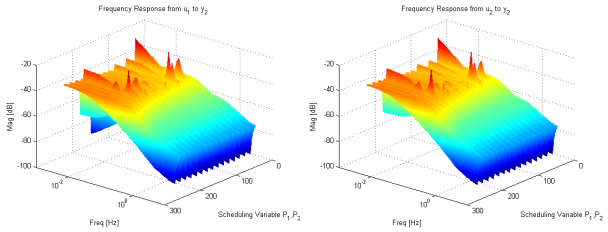
(d) From 2nd input to 2nd output

Figure A-3: Frequency response of LPV model with coefficient interpolation using bilinear interpolation method



(a) From 1st input to 1st output

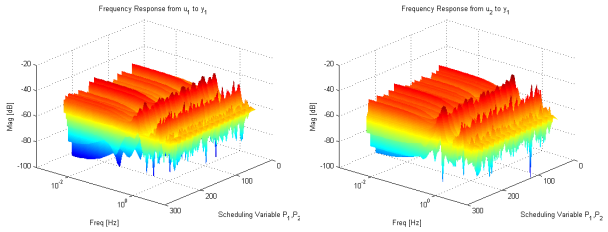
(b) From 2nd input to 1st output



(c) From 1st input to 2nd output

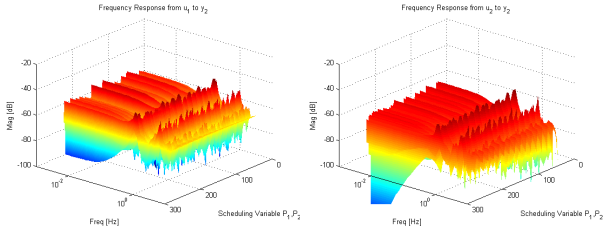
(d) From 2nd input to 2nd output

Figure A-4: Frequency response of LPV model with output interpolation with using RBF interpolation method



(a) From 1st input to 1st output

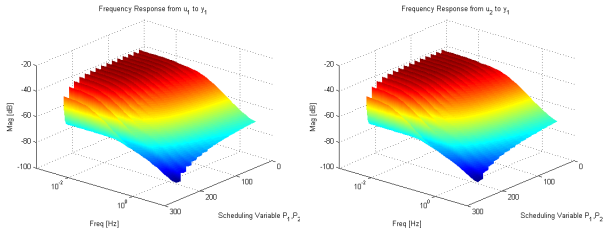
(b) From 2nd input to 1st output



(c) From 1st input to 2nd output

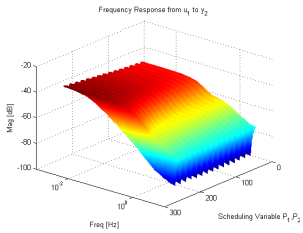
(d) From 2nd input to 2nd output

Figure A-5: Frequency response of LPV model with output interpolation using polynomial interpolation method

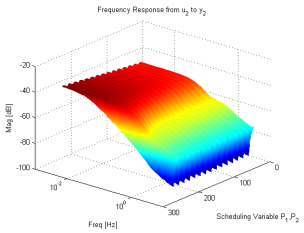


(a) From 1st input to 1st output

(b) From 2nd input to 1st output



(c) From 1st input to 2nd output



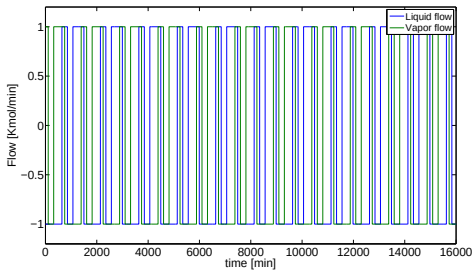
(d) From 2nd input to 2nd output

Figure A-6: Frequency response of LPV model with output interpolation using RBF interpolation method

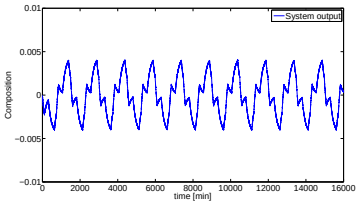
Appendix B

List of data set

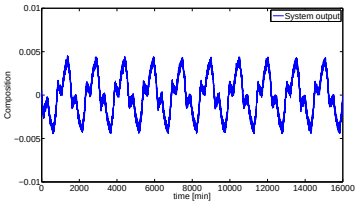
In this section example of data set used for identification($\mathcal{D}_I$), parameter optimization($\mathcal{D}_O$), and validation($\mathcal{D}_V$) can be found here.



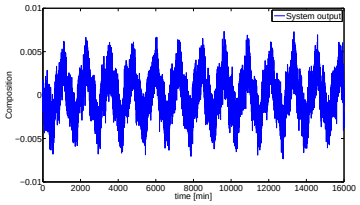
(a) Distillation column input on an operating point



(b) Output on low noise

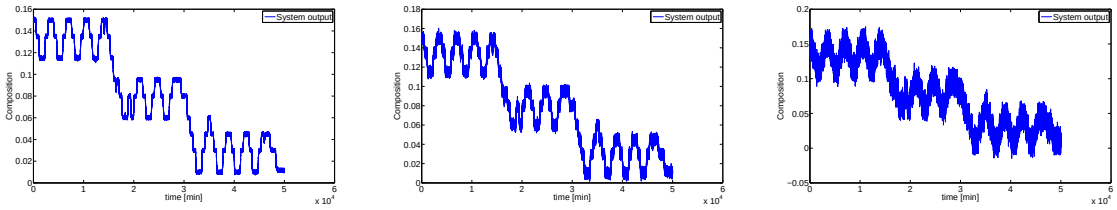


(c) Output on med noise



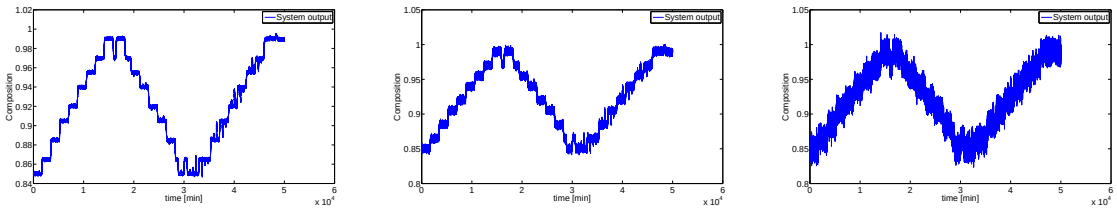
(d) Output on hi noise

Figure B-1: example of Input-Output data set used for identification of PEM model.



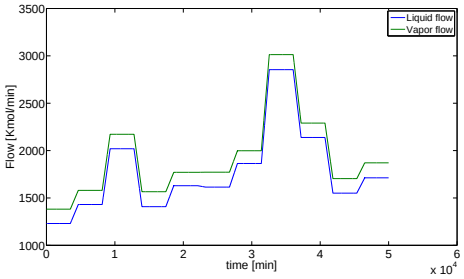
(a) Output on low noise **(b)** Output on med noise **(c)** Output on hi noise

Figure B-2: example of bottom product data set used for identification on global approaches.

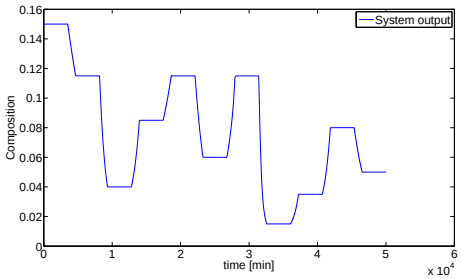


(a) Output on low noise **(b)** Output on med noise **(c)** Output on hi noise

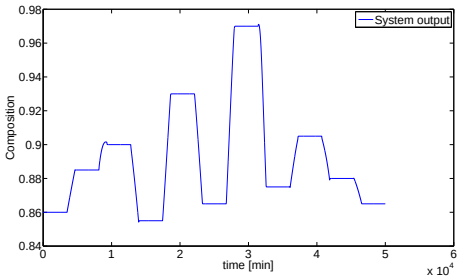
Figure B-3: example of top product data set used for identification on global approaches.



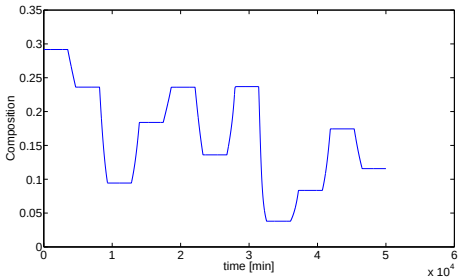
(a) Distillation column input



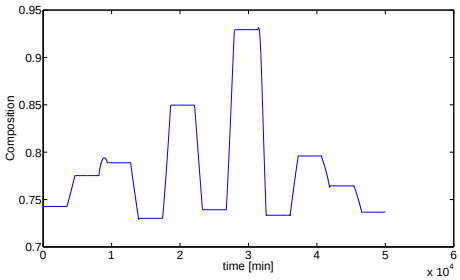
(b) Bottom product composition



(c) Top product composition

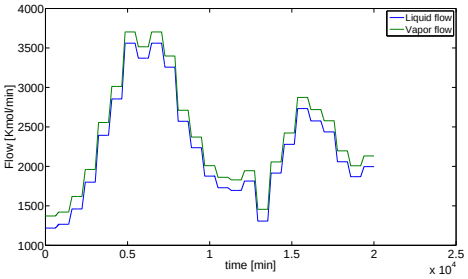


(d) 1st scheduling variable

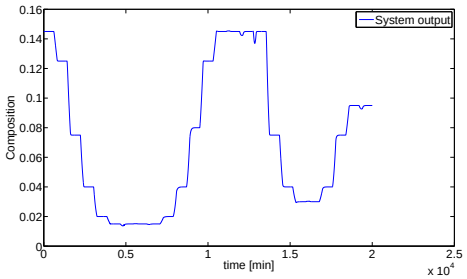


(e) 2nd scheduling variable

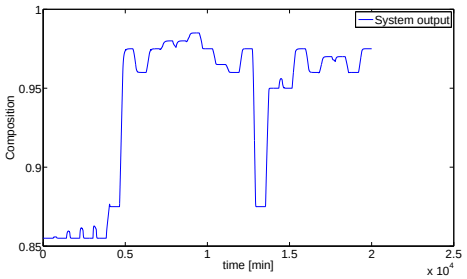
Figure B-4: Data set used for hyper parameter tuning in global approaches.



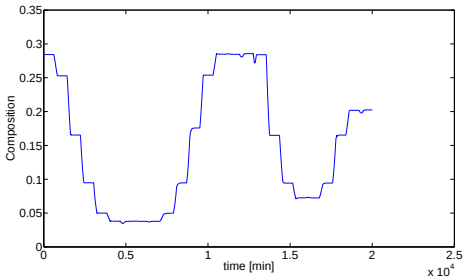
(a) Distillation column input



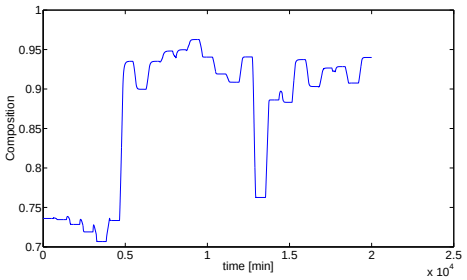
(b) Bottom product composition



(c) Top product composition



(d) 1st scheduling variable



(e) 2nd scheduling variable

Figure B-5: Data set used for validation.

Bibliography

[1] J. Shamma and M.Athans, "Gain scheduling: potential hazards and possible remedies," *IEEE Control Systems Magazine*, vol. 12, pp. 101–107, June 1992.

[2] A. Packard, "Gain scheduling via linear fractional transformations," *System Control Letters*, vol. 22, pp. 79–92, 1994.

[3] P. Apkarian and P.Gahinet, "A convex characterization of gain-scheduling $\mathcal{H}_\infty$ controllers," *IEEE Trans. on Automatic Control*, vol. 40, pp. 853–864, 1995.

[4] C. Scherer, "Mixed $\mathcal{H}_2/\mathcal{H}_\infty$ control for time-varying and linear parametrically-varying systems," *Int. Journal of Robust and Nonlinear Control*, vol. 6, pp. 929–952, 1996.

[5] K. Zhou and J.C.Doyle, *Essentials of Robust Control*. Prentice-Hall, 1998.

[6] Jiangyin Huang., Guoli ji, and Yucai Zhu, "Some study on the identification of multi-model lpv models with two scheduling variables," *Proceedings of the 11th IFAC Symposium on System Identification*, July 2012.

[7] M. Skogestad and M.Morari, "Control of ill-conditioned plants: High purity distillation," *AIChE Annual Meeting*, Nov 1986.

[8] M. Skogestad and M.Morari, "Control high-purity distillation columns from a robustness viewpoint," *Proceedings American Control Conference*, pp. 1325–1330, June 1987.

[9] M. Skogestad and M.Morari, "LV-Control of a high-purity distillation column," *Chemical Engineering Science*, vol. 43, pp. 33–48, 1988.

[10] M. Finco, W.L.Luyben, and R.E.Polleck, "Control of distillation columns with low relative volatilities," *Ind. Eng. Chem. Res*, vol. 28, pp. 75–83, 1989.

[11] M. Skogestad and M.Morari, "Control configuration selection for distillation columns," *AIChE, Journal*, vol. 33, pp. 1620–1635, 1987.

[12] M. Skogestad, P.Lundström, and E.W.Jacobsen, "Selecting the best distillation control configuration," *AIChE, Journal*, vol. 36, pp. 753–754, 1990.

[13] M. Skogestad, P.Lundström, and E.W.Jacobsen, "Modelling and identification for robust control of ill-conditioned plants - a distillation case study," *Proceedings of American Control Conference*, pp. 242–248, June 1991.

[14] I. Halvorsen and S.Skogestad, *Encyclopedia of Separation Science, Distillation.* Academic Press, Aug 2000.

[15] A. Huesman, "Process dynamics and control," *Lecture Notes in Delft University of Technology*, 2011.

[16] B. Ninness and F.Gustafsson, "A unifying construction of orthonormal bases for system identification," *IEEE Trans. on Automatic Control*, vol. 42, pp. 515–521, 1997.

[17] P.S.C. Heuberger, P.M.J. Van den Hof, and B. Wahlberg, *Modeling and Identification with Rational Orthonormal Basis Functions.* Springer-Verlag, 1997.

[18] R. Tóth, "Modeling and identification of linear parameter-varying systems," *Lecture Notes in Control and Information Sciences*, vol. 403, 2010.

[19] R. Tóth, P.S.C. Heuberger, and P.M.J. Van den Hof, "Orthonormal basis selection for LPV system identification, the Fuzzy-Kolmogorov c-Max approach," *Proceedings of the 45th IEEE Conf. on Decision and Control*, pp. 2529–2534, Dec 2006.

[20] J. Willems, "The behavioral approach to open and interconnected systems," *IEEE Control Syst. Mag.*, vol. 27, pp. 46–99, Dec 2007.

[21] R. Tóth, P.S.C. Heuberger, and P.M.J. Van den Hof, "The behavioral approach to linear parameter varying systems," *Automatic Control, IEEE Transactions*, vol. 56, pp. 2499–2514, Nov 2011.

[22] R. Tóth, "Modeling and identification of linear parameter-varying systems, an orthonormal basis function approach," *Ph.D.thesis, Delft University of Technology*, 2008.

[23] L. Lee and K.R.Poolla, "Identifiability issues for parameter-varying and multidimensional linear systems," *Proceedings of ASME Design, Engineering Technical Conference*, 1997.

[24] Y. Zhu and X. Xu, "A method of LPV model identification for control," *Proceedings of the 17th IFAC World Congress*, pp. 5018–5024, 2008.

[25] Y. Zhu and G. Ji, "LPV model identification using blended linear models with given weightings," *Proceedings of the 15th IFAC Symposium on System Identification*, pp. 1674–1679, 2009.

[26] R. Martínez-González, Bolea.Y., Grau.A., and H. Martínez-García, "an LPV fractional model for canal control," *Mathematical Problems in Engineering*, pp. 1–18, 2009.

[27] Bolea.Y., Blesa.J., and Puig.V., "LPV modeling and identification of an opel canal for control," *Proceedings of the European Control Conference*, pp. 3584–3591, 2007.

[28] B. Bamieh and L.Giarré, "Identification of linear parameter varying models," *Proceedings of the 38th IEEE conf. on Decision and Control*, pp. 1505–1510, Dec 1999.

[29] B. Bamieh and L. Giarré, "Identification for a general class of LPV models," *Proceedings of the 11th IFAC Symposium on System Identification*, June 2000.

[30] X. Wei and L. D. Re, "On persistent excitation for parameter estimation of quasi-LPV systems and its application in modeling of diesel engine torque," *Proceedings of the 14th IFAC Symposium on System Identification*, pp. 517–522, Mar 2006.

[31] Laurain.V., Gibson.M., Tóth.R., and Garnier.H., "Refined instrumental variable methods for identification of LPV Box-Jenkins models," *Automatica*, vol. 46, pp. 959–967, 2010.

[32] Tóth.R., Lyzell.C., Enqvist.M., P.S.C. Heuberger, and P.M.J. Van den Hof, "Order and structural dependence selection of LPV-ARX models using a nonnegative garrote approach," *Proceedings of the 48th IEEE Conference on Decision and Control*, pp. 7406–7411, Dec 2009.

[33] Leep.H. and Pötscher.B., "Sparse estimators and the oracle property, or the return of hodges' estimator," *Journal of Econometrics*, vol. 142, pp. 201–211, 2008.

[34] F. Previdi and M.Lovera, "Identification of a class of linear models with nonlinearly varying parameters," *Proceedings of the European Control Conf*, Aug 1999.

[35] F. Previdi and M.Lovera, "Identification of a class of nonlinear parametrically varying models," *Proceedings of the European Control Conf*, pp. 3086–3091, Sep 2001.

[36] F. Previdi and M.Lovera, "Identification of a class of nonlinear parametrically varying models," *Int. Journal on Adaptive Control and Signal Processing*, vol. 17, pp. 33–50, 2003.

[37] F. Previdi and M.Lovera, "Identification of nonlinear parametrically varying models using separable least squares," *Int. Journal of Control*, vol. 77, pp. 1382–1392, 2004.

[38] Hsu.K., Vincent.T.L., and Poolla.K., "Nonparametric methods for the identification of linear parameter varying systems," *Proceedings of the Int. Symposium on Computer-Aided Control System Design*, pp. 846–851, Sep 2008.

[39] L. Ljung, *System Identification, theory for the user*. Prentice-Hall, 1999.

[40] V. Vapnik, *Statistical Learning Theory*. Wiley-Interscience, 1998.

[41] J. Suykens and J.Vandewalle, "Least squares support vector machnine classifiers," *IEEE Trans. Circuits Syst. I Fundam Theory*, vol. 47, Nov 2000.

[42] V. Laurain, Tóth,R, and Wei Xing Zheng, "Introducing instrumental variables in the LS-SVM based identification framework," *Proceedings of the 50th IEEE Conf. on Decision and Control*, December 2011.