SOFT SENSOR DESIGN FOR DISTILLATION COLUMNS

Anna Maria Pulis
SOFT SENSOR DESIGN FOR
DISTILLATION COLUMNS

ANNA MARIA PULIS

Supervisors:
Professor Roberto Baratti
Professor Jesus Alvarez Calderon

Dottorato di Ricerca in Ingegneria Industriale
Università degli Studi di Cagliari
XIX Ciclo
Acknowledgements

I wish to thank Professor Roberto Baratti for his support and advice during my three years at the University of Cagliari.

I am also grateful to Professor Jesus Alvarez who was my supervisor both when I spent five months of 2005 at the Universidad Autonoma Metropolitana of Mexico City and also during my stay here at Cagliari.

Thanks also to the members of the process control group at the chemical department of Cagliari and at UAM. In particular I would like to thank Carlos Fernandez who I have shared many fruitful discussions.

The Fondazione Banco di Sardegna is kindly acknowledged for the support through its fellowship.
Contents

Abstract .............................................................................................................................. 1

Chapter 1. Introduction .................................................................................................. 3
  1.1 Introduction ............................................................................................................ 5
  1.2 Distillation Modeling ............................................................................................ 6
  1.3 Distillation Control ............................................................................................... 7
  1.4 State Estimation of Distillation Column: State of the art .................................. 8
  1.5 Thesis Objectives ............................................................................................... 11
  1.6 Thesis Organization ............................................................................................ 11
  1.7 Presentations ...................................................................................................... 12

Chapter 2. Distillation Columns ................................................................................... 15
  2.1 Introduction ......................................................................................................... 17
  2.2 The Balance Equations ....................................................................................... 18
  2.3 The Experimental Set-up .................................................................................... 20
    2.3.1 The pilot column .......................................................................................... 20
    2.3.2 Instrumentation ......................................................................................... 20
    2.3.3 Data sampling and control ....................................................................... 21
    2.3.4 The mixtures .............................................................................................. 22
    2.3.5 Dynamic test runs ..................................................................................... 23
  2.4 Thermodynamic Considerations ......................................................................... 25
  2.5 Model behavior ................................................................................................. 27

Chapter 3. Nonlinear State Estimation ........................................................................ 29
  3.1 Introduction ......................................................................................................... 31
  3.2 The Estimation Structure ..................................................................................... 32
    3.2.1 System reconstructibility .......................................................................... 33
    3.2.2 The temperature location criteria .............................................................. 38
    3.2.3 The temperature sensor number ............................................................... 40
  3.3 Algorithms .......................................................................................................... 41
    3.3.1 Innovation mechanism .............................................................................. 41
    3.3.1 The Kalman Filter ....................................................................................... 42
    3.3.2 The Extended Kalman Filter ..................................................................... 45
    3.3.3 The tuning ................................................................................................. 47
3.3.4 The Nonlinear Geometric Estimator (NGE) .......................................... 49

Soft Sensor Design ................................................................................................. 57
Introduction ......................................................................................................... 59

Part I: Estimation Structure Design ................................................................. 61

Chapter 4. The Data Assimilation Mechanism ................................................... 63
4.1 The Estimation Problem .............................................................................. 65
4.2 Adjustable-Structure Assessment ............................................................... 66
  4.2.1 Detectability structures ........................................................................ 66
  4.2.2 Detectability measures ......................................................................... 67
4.3 Structural Assessment ................................................................................ 68
  4.3.1 The single tray detectability analysis .................................................. 69
  4.3.2 Passivity concepts .............................................................................. 71
4.4 The Ternary Column ................................................................................. 73
  4.4.1 The coupled structure ....................................................................... 73
  4.4.1 Passive structures .............................................................................. 75
  4.4.2 The passivated structure ................................................................... 80
  4.4.3 Generalization .................................................................................. 81
4.5 Fixed versus adjustable structure .............................................................. 82

Part II: Structure Algorithm Testing ............................................................... 83

Chapter 5. The Soft Sensor Design ................................................................... 85
5.1 Introduction ................................................................................................. 86
5.2 The Soft Sensor Implementation ............................................................... 87
  5.2.1 The Nonlinear Geometric algorithm (NGE) ........................................ 87
  5.2.2 Construction and tuning .................................................................... 87
  5.2.3 The coupled Soft Sensor ................................................................... 88
  5.2.4 The single-sensor single-innovated state .......................................... 90
  5.2.5 The decoupled Soft Sensor ............................................................... 95
5.3 Extended Kalman Filter with reduced data injection ................................. 98
5.4 Comparison between GE, GEKF and CEKF ............................................ 99
5.5 The Adaptive-Adjustable Soft Sensor ........................................................ 100
5.6 Some Concluding Remarks ...................................................................... 102

Chapter 6. The Adaptive Kalman Filter .......................................................... 105
6.1 Introduction ............................................................................................... 107
6.2 The Covariance Matrices of the Estimates ................................................. 108
6.3 The a priori CEKF .................................................................................. 110
6.3.1 Results and discussion ................................................................. 112
6.3.2 Remarks ..................................................................................... 114

Chapter 7. Conclusions ..................................................................... 117

Bibliography ...................................................................................... 119
Abstract

In this Thesis, the issues concerning the jointly design of estimation structure and algorithm for distillation columns is addressed. By structure it is meant the number of measurements, their locations, the innovated-noninnovated state partition, and the innovated states per measurement. The algorithm is the dynamic data processor that performs the estimation task. The employment of a structure with partial innovation yields an estimator (soft sensor) with a simpler and more efficient data assimilation mechanism. The resulting estimator amounts to the combination of a low-order estimator for the innovated state and an open-loop estimator for the noninnovated state, and the possibility of adjusting the structure over a transient is also enabled. The Nonlinear Geometric (NGE) and the Conventional Extended Kalman Filter (CEKF) estimators were regarded as two algorithm design options. A tuning procedure for the CEKF and a reduce Kalman filter (GEKF) are proposed. The proposed approaches have been tested with the data generated by the transient response of a 32-stage (tert-butanol/ethanol/water) and (ethanol/water) pilot column, over an ample region of the state space. It was found that (i) the structural decisions plays a key role in the estimator behavior, regardless of the particular estimation algorithm employed (ii) the best estimator functioning is obtained by injecting the temperature measurement-based information over a few column states, and (iii) the resulting estimators are considerably simpler than the ones obtained with conventional geometric and CEKF techniques underline by a complete innovation structure, (iv) a systematic tuning procedure, based on an approximation of the model error covariance matrix, is useful to obtain a reliable estimator avoiding a time-consuming trial and error procedure.
Chapter 1

Introduction

In this Chapter the estimation problem in distillation columns are introduced, including the problem motivation, the related state of the art, the issues concerning the soft sensor design and the delimitation of the contribution of this investigation.
1.1 Introduction

Distillation is the most important unit operation in process industry where a given feed is separated in two or more products. The process separation requires large amounts of energy. In many plants it can involves 30-40% of the total investments and energy costs (Jacobsen, 1991). The energy requirements may be reduced by means of optimal column design or when the column is available with a system control able to maintain the optimal conditions.

A weak composition control can involve large periods of operation with purer products than necessary and consequently losses of energy (or products) that do not respect the specified desiderate values. This is the reason for which a robust column control is important to save energy and also to increase profit through better product recovery (Stichlmair, 1998; Kister, 1990).

The objective of a distillation column control scheme is to maintain the product composition close their desiderate values. Composition measurements are often expensive and difficult to maintain and also when their installation is possible they can introduce undesirable time-delays in the control loop. The lack or high costs of composition analyzers for industrial distillation columns motivates the design of temperature data driven soft sensors to infer on-line compositions on the basis of a process model and available temperature measurements (Stichlmair, 1998).

The soft sensor design involves the selection of the estimation structure and the choice of the estimation algorithm (i.e., the kind of dynamic data processor that performs the estimation task). The structural decisions (number and sensor location, set of states to estimate) play an essential role on the estimator implementation. The choice of number and sensor location still remains an open issue, with results that are not clearly connected with the estimator algorithm design and implementation (Roffel et al., 2003; Tronci et al., 2005; Zamprogna et al., 2005; Luyben 2005; Venkateswarlu and Kumar, 2006; Oisiovici and Cruz, 2000-2001).

The general consensus in the control community has been that only few temperatures should be used. The choice to consider the minimum number of measurements is often the most convenient with respect estimator design and tuning, also it avoids solving an over-parameterize problem. Some authors (c.f., Mejdell et al., 1991-1993) have put in discussion these statements, pointing out that the importance is not in the sensor number but in the relationship between measurements and output. Therefore the previous statements raise the following
questions: how many available measurements are really needed? The information content in measurements remains always significant or changes with time?

In literature is available a large variety of algorithms characterized by different structural design and tuning techniques. However, it is not clear the connection between structural property and algorithm selection. This implies the following questions: which algorithm can be used to develop the soft sensor? Can the structural property compromise the algorithm selection?

The above issues are the focal point of this PhD Thesis which objective is to provide guidelines to jointly design the estimation structures and algorithms in binary and multicomponent columns.

In the next three Sections will be discussed the state of the art related to the present contribution, and then the Thesis objectives will be stated.

1.2 Distillation Modeling

In order to design a robust soft sensor for a distillation column is necessary to understand the column behavior and this could be obtained through experiments on operating columns or by means of a mathematical model of the process.

The behavior of distillation columns has been extensively studied over the last years and still remains an active area of research. The industrial researchers Rosenbrock (1962) and Radamaker (1975) can be considered the pioneers in dynamic modeling and simulation, and on understanding column behavior (Skogestad, 1992; Kister, 1990).

A large variety of model for distillation column has been proposed: steady state or dynamical model, simplified or rigorous. Where the term “rigorous” is usually referred to a staged model that includes mass, energy balances on each stage, liquid flow dynamics (changes in liquid hold-up) and pressure dynamics on each stage.

In the literature three different “full order” models, which incorporate differential equations for balance on enthalpy (E), mass (M), and chemical species (C) on each plate, are commonly used. These are sometimes denoted: (i) EMC, (ii) MC, and (iii) C (Levy et al., 1969; Skogestad, 1992). The EMC model is a rigorous model with the energy balance included, the assumptions of negligible vapor hold-up or constant pressure is used. The MC model denotes a model with negligible vapor hold-up, constant pressure and constant molar flows such that the energy balance is not needed. In conclusion, the C model corresponds to a model where constant
liquid hold-up is assumed such that the overall material balance is not needed (Skogestad, 1992).

For control or estimation purposes frequently are adopted simplified models based on the common assumptions: a) neglecting hold-up in the vapor phase, b) constant pressure, c) equal vapor flows up the column. Skogestad (1992) suggested not using a model with constant liquid dynamic hold-up for control purposes. The importance of the liquid dynamic hold-up to monitor and control was also underlined by Fernandez (1996) and Castellanos-Sahagun (2005).

There are of course a lot of combinations and simplifications possible in addition to the ones above. For example, it is common assumes an ideal VLE. However, this assumption is not possible when the vapor-liquid equilibrium has a strong nonlinear behavior (Baratti et al., 1998; Oisiovici et al., 2000) because the values selected for the interaction coefficient can strongly influence the entire simulation.

Without restricting the approach, the estimation model adopted in this Thesis is simple and characterized by the standard assumption above mentioned, i.e., the liquid dynamic hold-up is considered in quasi steady state regime but the molar tray hold-up is update in time using the Francis formula. This allows having a hold-up updated during the simulation and in the same time to have a less computational burden (the overall mass balance is neglected).

It is important to remark that, being the objective of this Thesis to design robust estimators able to monitoring on-line product composition, the model adopted is simple but able to capture the main features of the column dynamics.

1.3 Distillation Control

Given that control is a motivation of estimator design, in this Section some words on the matter are said. The major objective of any control system for any process is the maintenance of its operation about a prescribed output, without any human action (Stichlmair and Fair, 1998). This underlines the necessity to avoid or to eliminate disturbances caused from exogenous input changes. Sometimes, a process control system can consists of a large number of control loops that may interact each other. A simple example of an interacting system is a distillation column with dual composition control.

Generally, in a distillation column the variables to be regulated are the product compositions in the bottom and in the distillate. Usually, a distillation column may be regulated by a 5×5 control system, with 5 manipulated variables and 5 controlled variables. An optimal controller should manipulate all five inputs to keep all the
outputs close to the desiderate values. The problem of the column control could be solved by a multivariable $5 \times 5$ controller taking in account all possible interaction among the variables. In reality, few columns are controlled using a full $5 \times 5$ controller. The common practice to control product composition is a decentralized system with single loop controllers (Jacobsen, 1991; Luyben, 1990; Skogestad et al., 1990; Skogestad et al., 1988; Skogestad and Morari, 1988; Stichlmair and Fair, 1998).

When a decentralized control structure is adopted then the control design can be divided into: (i) Choose the inputs and design level and pressure control, (ii) Design composition controller (Jacobsen, 1991). Guidelines to design level and pressure controller are also provided in the literature (Shinskey, 1984, Luyben, 1990), while the decision if a full $2 \times 2$ controller or a single loop controller should be used for composition control is not straightforward and generally are used single loop controllers also for composition control.

This means to use five SISO controllers and consequently the question is how to pair the input variables with the output variables. A large variety of different control configurations for distillation column have been developed. The selection of control configuration is an important issue that is extensively discussed in literature (Luyben, 1970; Stichlmair 1998; Lundstrom and Skogestad, 1995; Waller et al., 1988; Skogestad and Morari, 1988; Castellanos-Sahagun et al., 2005). On the other hand, there is not single configuration, which is suitable for all columns. This has been noted already by Boyd (1946) who states: “It would be impossible to give a control system that would be a panacea for all fractionation control problem”.

In this Thesis, all the experimental runs considered were carried out by manipulating the distillate and bottom flows (proportional-only level controller), while the column was operated in open loop model by maintaining the feed, the steam to the reboiler and the reflux at the desiderate values.

### 1.4 State Estimation of Distillation Column: State of the art

A key problem in chemical process is to control, in a reliable and inexpensive way, the dynamic evolution of the system states (i.e. product compositions, reactants, temperatures).

In many processes only few product composition measurements are available online to control the product quality. This is due to the absence or high investment and maintenance costs of composition analyzers that often, when available, introduce undesirable time delays in the control loop.
These problems can be overcome adopting an inferential system or a soft sensor that permits to reconstruct the system dynamic by using a process model and available on-line measurements.

Different works are addressed to design soft sensor for distillation columns and the state of the art can be seen in Joseph and Brosilow (1978), Gilles and Retzbach, (1983); Lang and Gilles, 1990; Quintero-Marmol et al. (1991), Baratti et al. (1995, 1998), Mejdoll and Skogestad (1991, 1993), Oisiovi and Cruz (2000, 2001), Tronci et al. (2005), Zamprogna et al. (2005), Venkateswarlu and Kumar (2006). Specifically, in state estimation area, many works are devoted to estimate product composition in batch and continuous distillation columns by means of soft sensors. Brosilow (1978) was considered the pioneer in the development of optimal and sub-optimal estimators to infer product composition in a continuous distillation column by means of temperature and flow measurements. Since then, several authors have studied this topic.

Yu and Luyben (1987) designed the nonlinear Brosilow’s estimator for multicomponent systems. They analyzed the main issues in state estimator design and stated that a distillation column to be observable, at the steady state condition, needed a temperature sensor number equal to (Nc -1), where Nc is the mixture component number.

Lang and Gilles (1990) presented a full-order nonlinear observer, based on a simplified model, for binary (methanol/water) and ternary (methanol/ethanol/n-propanol) columns. The concept of Zeitz (Isidori, 1985) has been applied to design the observer that is tested by simulation on a 40-tray column. They studied the column observability in order to individuate the column region characterized by an appropriate temperature gradient. The advantages of the observer proposed by Lang and Gilles are its very simple structure and the fact that only few parameters must be tuned.

Quintero-Marmol et al. (1991) applied and extended Luenberger observer (ELO) to predict compositions in multicomponent batch distillation columns from temperature measurements. The observer is a deterministic estimator, which may not perform well in the presence of both process model mismatch and noise in the measurements.

Deza and Gauthier (1992) presented full order nonlinear observers for distillation columns. The error of these observers converges exponentially to zero. The methodology consists in studying, first, the system observability and in applying a nonlinear coordinate change in order to have the model in a canonical form (the model is constituted by a linear and a nonlinear block). By means of a coordinate change is obtained a triangular structure of the system to which was applied an EKF.
Mejdell and Skogestad (1991, 1993) discussed the use of the static principal-component-regression (PCR) and partial-last-squares (PLS) estimators and implemented them on a pilot-plant column. Moreover, they compared the performance of the Kalman Bucy Filter, Brosilow’s inferential estimator and the PCR underlying the importance of the information content in measurements.

Baratti et al. (1995, 1998) applied a nonlinear Extended Kalman filter to infer the compositions of the streams leaving a binary (ethanol/water) and a ternary (ethanol/tert-butanol/water) continuous distillation column from temperature measurements. The estimator performance was compared with the dynamic behavior of an experimental 30-trays column. The authors observed that accurate description of vapor liquid equilibrium is important in continuous distillation columns to achieve a good estimator performance.


Tronci et al. (2005) presented a low order NGE (Nonlinear Geometric Estimator), based on the differential geometry, to infer product composition in a binary column where it is separated the mixture ethanol-water.

As it is possible to observe, the distillation monitoring problems have been extensively tested with a diversity of techniques. It is important to point out that the nonlinear extended Kalman filter (EKF) is the most widely used estimation technique, but it is characterized by an implementation that requires the tuning of covariance via trial-and-error and the on-line integration of a set of auxiliary ordinary differential equations (ODEs) whose number grows rapidly with stage number and components.

Only in few works the estimator design is addressed by means of an accurate analysis of the structural observability property of the system. This implies the need to gain more general insight by means of a better understanding of the system and of the information content in measurements.

In this Thesis, considering the column models developed by Baratti et al. (1998), the evolution of the states dynamic in binary and ternary systems by means of secondary temperature measurements is reconstructed. In particular, the soft sensor design is developed by means of a methodology, based on a geometric approach, which permits to select the estimation structure better suited for the column studied.
1.5 Thesis Objectives

The main objective of this work is to develop a methodology to jointly design the estimation structure and algorithms, with binary and ternary system as case example with experimental data.

The methodology adopted is based on:

(i) Studying the data assimilation mechanism along the column in order to have information concerning the column region with a significative relationship between the input measurements and the output to estimate;

(ii) Choosing an adaptive-adjustable estimation structure (number and sensor location, set of the states to estimate) suitable for each algorithm;

(iii) Designing a complete or reduced algorithm with the possibility of changing the estimation structure during the column operation were considered;

(iv) Proposing a tuning procedure based on the approximation of the covariance matrices and on physical consideration, alternative to the trial and error technique.

The contribution of the present work consists in propose a procedures to select robust estimation structures and a systematic tuning technique.

1.6 Thesis Organization

This Thesis deals with the design of soft sensors for distillation columns. The Thesis contains three principal parts: the first part deals with estimation structure selection by means of a methodology based on a geometric approach; the second part deals with the design of soft sensor based on the candidate structure selected; while the third part concerns with the proposal of a tuning procedure alternative to the trial and error tuning technique.

A brief overview of each Chapter is given below:

In the Chapter 2 the column model, the pilot plant and the thermodynamic analysis are illustrated.

In the Chapter 3 the nonlinear estimation tools and the algorithms adopted to design the soft sensor are presented.

In the Chapter 4 the methodology used to select the estimation structure based on the differential geometry is developed.
In the Chapter 5 the jointly design of the estimation structure and algorithms is presented.

In the Chapter 6 a tuning procedure based on approximation of the covariance matrices is proposed.

In the Chapter 7 the conclusions are illustrated.

1.7 Presentations

Preliminary versions of some of the topic presented in this Thesis have been presented and submitted to some international and national chemical engineering conferences:

Chapter 4, 5: International Symposium on Advanced Control of Chemical Processes (ADCHEM-IFAC), Gramado (Brasile), 2-5 April 2006.


Accepted to the 8-th International Symposium on Dynamic and Control of Process System, Cancun (Mexico), 04-06 June 2007

Chapter 6: Conference on Chemical and Process Engineering (Icheap-7), Giardini di Naxos (Me), 15-18 June 2005

In addition to the work presented in this Thesis the author has also participated, during the course of study, in other research projects concerning the monitoring and controlling of wastewater treatment plants and the composition estimation in binary distillation columns:


Chapter 2

Distillation Columns

*In this Chapter the column model adopted to simulate the separation of a ternary mixture and the experimental investigations carried out on a pilot plant distillation column, located at the University of Padova (Italy), are presented. The physical property of the mixtures fractionated, the column operating conditions and a thermodynamic analysis of the system are described. The main objective of this Chapter is to have a reasonable model-based description of the column behavior, including validation with experimental data, or equivalently, the development of the model to be used in our structure-algorithm estimation study. In addition, some (rather drastic) test column evolutions are considered in order to test the estimation schemes.*
2.1 Introduction

Distillation is used to separate a binary or a multi-component feed stream in two or more products with different compositions. The physical principle on which is based the separation process is the difference in volatility of the components that must be separated. The separation takes place in a column where heat is added to the reboiler and removed from the condenser. The vapor stream rises (from the reboiler) through the column and it comes into a contact with the liquid stream rising from the reflux drum. The volatile component is enriched in the vapor phase and the heavy components are enriched in the liquid phase (Lundstrom, 1994; Kister, 1990).

Dynamic mathematical models are widely used in distillation column process design and operation, and the state of the art can be seen elsewhere (e.g. Luyben, 1990; Gani et al., 1986; Cho and Luyben, 1987). They have demonstrated to be powerful tools for a large number of researches: (i) to assess controller structures and to tune controllers; (ii) to evaluate the effects of disturbances; (iii) to optimize the plant operation and to investigate start-up and shutdown techniques.

In general, mathematical models for distillation columns are of high order and strongly nonlinear. This is due to the high number of stages and components that characterize an industrial column. The dynamic simulation of distillation columns involves numerical integration of a large set of ordinary nonlinear differential equation that arises from the material and energy balance on each tray. In addition, the separated time constants of the column give a stiff character to the system. For this reason it is important to know a priori the stiffness ratio and the system parameters that affect this ratio. Several works have dealt with the problem to correlate the physical parameters on the stiffness of the system (Lagar, 1987; Tyreus et al., 1975).

There has been a continuous development of distillation column models, with emphasis on process control, for maintaining product specifications. In this Chapter the column model adopted to simulate the fractioning of a ternary mixture is presented. In the first Section of the Chapter the mathematical equations and the assumption made to simulate the ternary column are briefly described, while in the second part the experimental set-up and a thermodynamic system analysis is introduced.
2.2 The Balance Equations

In order to simulate a distillation column, different models with different complexity levels (Gani et al., 1986) can be used. In this Section a model able to simulate the column behavior is presented. As already stated in the previous Section, the estimator must be used for on-line application, hence a simplified column model is adopted, and it is integrated by using the simple and fast Heun integration scheme.

The model adopted is the same presented by Baratti et al. (1998) and is based on the following assumptions:

(i) The energy balance on each tray is neglected, and the vapor and liquid flow rates are assumed constant in each section of the column. This choice is particularly suitable for the specific mixture under examination because the molar heat of vaporization for the two alcohols changes very little with compositions.

(ii) The dynamics of the tray molar hold-up is neglected, while its value is continuously updated in time by means of the Francis formula.

(iii) A linear pressure drop was assumed along the column, considering bottom and top pressure equal to 814 and 760 mmHg, respectively.

(iv) The reboiler has been modeled as an ideal separation stage.

In this way the adopted model accounts for the composition dynamics, while it does not describe the fast hydraulic of the column.

Consider a continuous \( N \)-stage ternary distillation column (Figure 2.1) with feed rate \( F \) at composition \( c_F \), distillate (\( D \)) and bottoms (\( B \)) at composition \( c_D \) or \( c_B \), heat load \( Q \) (proportional to vapor flow rate \( V \)), and reflux flow rate \( R \). The column dynamics are described by the following equations (Skogestad et al., 1997, Baratti et al., 1998):

Reboiler

\[
\frac{c_{i,0}}{\eta_0} = \frac{[R \Delta^+ c_{i,0} - V \Delta c_{i,0}]}{\eta_0}
\]

(2.1)

Stripping section \((1 \leq i \leq n_F, k = 1, 2)\)

\[
\frac{c_{i}^{k}}{\eta_i} = \frac{[(R + F) \Delta^+ c_{i}^{k} - V \Delta c_{i}^{k}]}{\eta_i}
\]

(2.2)
CHAPTER 2. DISTILLATION COLUMNS

Feed tray \((i = n_F, k = 1, 2)\)

\[
\begin{align*}
\hat{c}_{i,n_F}^k &= \left[ R \Delta^* \hat{c}_{i,n_F}^k \cdot V \Delta \hat{v}_{i}^k (c_{1,n_F}^k, c_{2,n_F}^k) + F (c_{i,F}^k - c_{i,n_F}^k) \right] / \eta_i
\end{align*}
\] (2.3)

Enriching section \((n_F + 1 \leq i \leq N-1, k = 1, 2)\)

\[
\hat{c}_{i}^k = \left[ R \Delta^* \hat{c}_{i}^k - V \Delta \hat{v}_{i}^k (c_{1,i}^k, c_{2,i}^k) \right] / \eta_i
\] (2.4)

Top Tray \((i = N, k = 1, 2)\)

\[
\hat{c}_{N}^k = \left[ R \Delta^* \hat{c}_{N}^k - V \Delta \hat{v}_{N}^k (c_{1,N}^k, c_{2,N}^k) \right] / \eta_i
\] (2.5)

Condenser

\[
\hat{c}_{D}^k = \left[ V (\hat{v}_{D}^k (c_{N}^k) - c_{D}^k) \right] / \eta_i
\] (2.6)

Temperature measurements

\[
\begin{align*}
{y}_s &= T_s = \beta(c_{1,s}^k, c_{2,s}^k), \ s \in [1, n_F - 1] \\
{y}_e &= T_e = \beta(c_{1,e}^k, c_{2,e}^k), \ e \in [n_F + 1, N]
\end{align*}
\] (2.7, 2.8)

where \((k = 1, 2)\)

\[
\begin{align*}
c_{1,i}^k + c_{2,i}^k + c_{3,i}^k &= 1, \quad \Delta^* \hat{c}_{i}^k = c_{i+1}^k - c_{i}^k
\end{align*}
\] (2.9)

\[
\Delta \hat{v}_{i}^k (c_{1,i}^k, c_{2,i}^k) = \hat{v}_{i}^k (c_{1,i}^k, c_{2,i}^k) - \hat{v}_{i-1}^k (c_{1,i-1}^k, c_{2,i-1}^k)
\] (2.10)

\(c_{1,i}^k\) (or \(c_{2,i}^k\)) is the 1st (or 2nd) component (molar fraction) composition in the \(i\)-th stage, \(y_s\) (or \(y_e\)) is the temperature \(T_s\) (or \(T_e\)) in the \(s\)-(or \(e\))-th stage (to be determined) of the stripping (enriching) section, \(\nu_i\) (or \(\nu_e\)) is the nonlinear (liquid-vapor equilibrium) function that determines the \(i\)-th component composition in the vapor phase, \(\beta\) is the nonlinear function that sets the temperature (bubble point), and \(\eta\) is the (tray hydraulics) function that sets the exit molar flow rate from the \(i\)-th stage.

Generally, the hold-up dynamic is faster than composition dynamic, and it is assumed in quasi-static regime. The molar tray hold-up is updated in time using the geometric relationship:

\[
\eta_i = \rho h_i \delta
\] (2.11)
where $\rho_i$ is the molar liquid density, $S$ is the active surface of the tray, $h_i$ is the liquid height, and $\delta$ accounts of the effect of the froth on the trays as a function of the vapor velocity. In the case of both the reboiler and the condenser, the liquid hold-up has been set equal to the value imposed by the level controller 0.05 and 0.012 m$^3$, respectively.

![Distillation Column](image)

Figura 2.1. Distillation Column.

### 2.3 The Experimental Set-up

The experimental apparatus is the same as the one used by Baratti et al. (1995, 1998). In this Section a brief description of the experimental pilot plant is given, more details can be seen in Da Rold (1992).

#### 2.3.1 The pilot column

The continuous distillation column (Figure 2.2) is about 10 meters high, has a diameter of 0.30 m, consists of 30 sieve-trays and the space between the trays is 0.2 m. The reboiler is the shell-and tube natural circulation termosyphon type (heat transfer area 4.4 m$^2$) fed with steam at $P = 202$ kPa. The overhead vapor is sub-cooled in a water-cooled condenser (heat transfer area 7.7 m$^2$) and the reflux drum is open to the atmosphere. The feed is introduced in the 8-th tray from the bottom.

#### 2.3.2 Instrumentation

The plant, sketched in Figure 2.2, is equipped with electro-pneumatic valves: feed, reflux, distillate, and bottoms streams. Several variables are measured and monitored on-line such as:

- Temperatures are measured by Pt-100 $\Omega$ resistance thermometers on trays 0, 4, 8, 12, 18, 22, 26, 30 from the bottom, in the feed and in the reflux stream.
— Volumetric flow rates (feed, reflux, distillate, bottom, and steam) are measured by turbine meters.
— Pressure differential transducers provide values of pressure in the reflux drum and in the column bottom.
— Simultaneous acquisition of temperature, flow rate and pressure of the steam allows calculating the reboiler duty and the internal vapor flow-rate at the column bottom.
— Distillate and bottom compositions were sampled every 3-5 minutes and measured off-line by means of a gas-chromatograph equipped by a resistance detector.

The column was operating in open loop; a volumetric flow control was used for both the feed and the reflux streams, a mass flow for the steam to the reboiler, while the top and bottom levels were regulated by the distillate and residue product.

![Figure 2.2. Distillation Column Pilot Plant.](image)

### 2.3.3 Data sampling and control

All analogue signals are conditioned, converted to digital ones, and transmitted to a PC via the standard RS-232 interface. A D/A converter is used to send the control signals from the computer to the plant through the same interface. The column is fully operated by the PC. A computer program was developed (Da Rold, 1992) whose main features are:

— Data acquisition, memorization and monitoring on different screens;
— Valve actuation for control of levels, reflux and feed volumetric flow rates;
— Control of the mass flow-rate of the steam fed to the reboiler;
— Safety checks (top and bottom levels, pressure drop, coolant flow-rate);
— Routine for driving the plant to the desired steady state;
— Possibility of superimposing disturbances to reflux, distillate, feed volumetric flow rates and steam mass flow rate.
2.3.4 The mixtures

The systems considered for the experimental investigations were the binary mixture ethanol/water and the ternary mixture ethanol/tert-butanol/water. Vapor-liquid equilibrium (VLE) on each tray is calculated by using the NRTL model for the computation of activity coefficient and assuming an ideal vapor phase. The composition of the vapor phase leaving the generic tray is calculated from the corresponding equilibrium value by using the Murphree efficiency equation. The activity coefficients for the binary and ternary mixture have been calculated with the parameters values; see Table 2.1, reported by Gmehling and Onken (1977a, b). From the same reference have also been taken the values of the Antoine parameters; see Table 2.2, for calculating the vapor pressure as a function of temperature.

Table 2.1 NRTL Parameters for the three binary mixtures.

<table>
<thead>
<tr>
<th></th>
<th>$G_{12}$</th>
<th>$G_{21}$</th>
<th>$\alpha$</th>
</tr>
</thead>
<tbody>
<tr>
<td>ethanol/tert-butanol</td>
<td>303.4404</td>
<td>-323.1464</td>
<td>0.3033</td>
</tr>
<tr>
<td>ethanol/water</td>
<td>363.6962</td>
<td>1181.1166</td>
<td>0.719</td>
</tr>
<tr>
<td>tert-butanol/water</td>
<td>1651.34</td>
<td>387.34</td>
<td>0.458</td>
</tr>
</tbody>
</table>

In the case of the ternary mixture the parameters values selected were obtained from binary data of the three binary mixtures ethanol/water, tert-butanol/water and ethanol/tert-butanol.

Table 2.2. Antoine constants P [mmHg] and T [°C].

<table>
<thead>
<tr>
<th></th>
<th>A</th>
<th>B</th>
<th>C</th>
</tr>
</thead>
<tbody>
<tr>
<td>ethanol</td>
<td>8.11220</td>
<td>1592.864</td>
<td>226.184</td>
</tr>
<tr>
<td>tert-butanol</td>
<td>7.36168</td>
<td>1180.930</td>
<td>180.476</td>
</tr>
<tr>
<td>water</td>
<td>8.07131</td>
<td>1730.630</td>
<td>233.426</td>
</tr>
</tbody>
</table>

It is important to note that the equilibrium ethanol/water is very well known in all its composition range, while for tert-butanol/water only few data are available mostly in the middle and high-range of alcohol composition. The third system ethanol/tert-butanol is characterized by an almost ideal behavior.

Table 2.3 Azeotrope conditions at atmospheric pressure.

<table>
<thead>
<tr>
<th></th>
<th>B.P [°C]</th>
<th>Azeotropic conditions</th>
</tr>
</thead>
<tbody>
<tr>
<td>ethanol/water</td>
<td>78.3-100</td>
<td>$x_{\text{alcohol}} - x_{\text{w}}$ mol%</td>
</tr>
<tr>
<td>tert-butanol/water</td>
<td>82-100</td>
<td>0.35 - 0.65</td>
</tr>
</tbody>
</table>
Both the systems ethanol/water and the tert-butanol/water present azeotrope at the conditions reported on Table 2.3 and at atmospheric pressure.

### 2.3.5 Dynamic test runs

The experimental data considered to test the estimator performance are the same presented by Baratti et al. (1995, 1998). However, the experimental procedure adopted in each experimental run can be summarized as follows. First the column operating conditions were selected in order to obtain the desiderate steady state, using a steady state simulation program. Next, the control program was activated to start up the plant. Steady state conditions were usually reached in less than 1 hour. At this point, the scheduled changes in the operating conditions were applied to the column and the resulting behavior was monitored by measuring on line the temperature on selected trays and by sampling distillate and bottom streams.

Several experiments were carried out, but in particular were considered those that starting from a steady state with low alcohol end up with high alcohol content in the distillate. In Tables 2.4 and 2.5 are represented the operating conditions of the experimental runs carried out with the ternary and binary mixtures, respectively. All the experiments are characterized by: (i) constant feed composition, (ii) a sub-cooled temperature feed ($T_F \sim 20-25 \, ^\circ C$), and (iii) a sub-cooled reflux stream. In the first test (Run I) the column transient was induced decreasing the vapor flow-rate, while in the second and third tests (Run II, Run III) increasing the reflux and feed flow rates, respectively.

<table>
<thead>
<tr>
<th>Run</th>
<th>Run I</th>
<th>Run II</th>
<th>Run III</th>
</tr>
</thead>
<tbody>
<tr>
<td>$R$ (mol/s)</td>
<td>0.77</td>
<td>0.58</td>
<td>0.7214</td>
</tr>
<tr>
<td>$F$ (mol/s)</td>
<td>1.323</td>
<td>1.434</td>
<td>1.3865</td>
</tr>
<tr>
<td>$(c_{E,T}^E, c_{T,T}^T)$</td>
<td>(0.0979, 0.0630)</td>
<td>(0.0835, 0.0535)</td>
<td>(0.0829, 0.0532)</td>
</tr>
<tr>
<td>$V$ (mol/s)</td>
<td>1.57</td>
<td>1.25</td>
<td>1.435</td>
</tr>
<tr>
<td>$\Delta R$</td>
<td>~ + 40%</td>
<td>~ + 35%</td>
<td>~ + 35%</td>
</tr>
<tr>
<td>$\Delta V$</td>
<td>~ + 35%</td>
<td>~ + 35%</td>
<td>~ + 35%</td>
</tr>
<tr>
<td>$\Delta F$</td>
<td>~ + 35%</td>
<td>~ + 35%</td>
<td>~ + 35%</td>
</tr>
</tbody>
</table>

The Figure 2.3 shows the evolution of the experimental ethanol and tert-butanol compositions for the Run I (a), Run II (b) and Run III (c) referred to the experiments carried out with the ternary mixture ethanol/tert-butanol/water.
Table 2.5 Column operating conditions for the binary mixture.

<table>
<thead>
<tr>
<th>Run</th>
<th>Run I</th>
<th>Run II</th>
</tr>
</thead>
<tbody>
<tr>
<td>R (m³/s)</td>
<td>3.4×10⁻⁵</td>
<td>3.4×10⁻⁵</td>
</tr>
<tr>
<td>F (mol/s)</td>
<td>1.78</td>
<td>1.83</td>
</tr>
<tr>
<td>(c^E_n)</td>
<td>0.140</td>
<td>0.137</td>
</tr>
<tr>
<td>V (mol/s)</td>
<td>1.45</td>
<td>1.56</td>
</tr>
<tr>
<td>ΔR</td>
<td>~ + 40%</td>
<td></td>
</tr>
<tr>
<td>ΔV</td>
<td>~ - 35%</td>
<td></td>
</tr>
</tbody>
</table>

Figure 2.3.a,b,c Evolution of ethanol and tert-butanol compositions in the Run I, Run II and Run III related to the ternary mixture.

For the ternary mixture, it is important to observe that in the Run III, given in Table 2.4, the alcohol composition at the bottom remains always below the detection limit. For this reason only Run I and II are considered to test the soft sensor performance. Moreover, the operating condition adopted and the step changes in the feed, reflux and vapor flowrate have been selected in order to have significative changes of top and bottom compositions. In the light of these considerations the “atypical” runs represent a severe test to design of state estimators for continuous and batch distillation columns. In particular, in this Thesis the experimental runs related to the
ternary mixtures have been used to evaluate the soft sensor robustness in Chapter 4 and 5, while the experimental runs related to the binary mixtures are considered in Chapter 6 to test the soft sensor performance with a tuning procedure alternative to the trial and error technique.

### 2.4 Thermodynamic Considerations

In order to design a robust soft sensor it is necessary to understand the column behavior. Therefore in this Section some thermodynamic considerations, useful for the soft sensor design purposes, are made. In particular, let us consider the column behavior when the ternary mixture ethanol/tert-butanol/water is fractionated. In the previous Section it was mentioned that eight temperature measurements are available, while composition measurements are not available on-line. For this reason in Figure 2.4 are shown the transients of the simulated temperature and alcohol compositions stage profiles (a, b and c) referred to Run I. The analysis of Figure 2.4 emphasizes that temperature dynamics are faster than composition dynamics. In fact, the composition profiles are still evolving in the enriching section, while the temperature transients reach steady state condition. This confirms that in a multicomponent mixture, the relationship between temperature and composition is not uniquely defined as in the binary case.

![Temperature and composition transient](image.png)

Figure 2.4.a,b,c. Temperature and composition transient.
As already stated in the previous Section, the experimental runs considered are those for which the column works at high purity condition. In the Figures 2.4(a) referred to Run I, it is possible to note that the temperature profile becomes rather flat in the enriching section (stages: 30, 26, 23, 18) after 20-30 minutes that the process separation is started. This is due to an oversize of the column for the mixture considered; indeed most of the stages in the enriching section work at very high purity conditions. Moreover, during the separation process the column works close to the azeotrope condition. Due to this fact, the temperature changes are significant only in the stripping section (low alcohols concentrations) while vanish in the enriching section (high alcohols concentrations) involving the presence of a flat region in the temperature surface, see Figure 2.5, where the temperature is almost constant ($T \approx 353$ K).

![Figure 2.5. Temperature surface.](image)

In particular, in the Figure 2.6 is shown the projection of the column distillation line evolution upon the temperature contour.

![Figure 2.6. Composition profiles at three different times.](image)
The distillation lines evolution, at three different times, start from low \((E_0^c \approx 0, T_0^c \approx 0, E_D^c \approx 0.3, T_D^c \approx 0.19)\) toward high alcohol compositions \((E_0^c \approx 0.019, T_0^c \approx 0.0043, E_D^c \approx 0.43, T_D^c \approx 0.31)\) moving close to a region where a small temperature change implies large composition variations. This suggests that the temperature-composition relationship is not uniquely defined and the estimation problem could become more difficult in the enriching section.

### 2.5 Model Behavior

Before considering the soft sensor design it is convenient to discuss the comparison between simulation results and experimental data referred to the Run I and II (Table 2.4). In the Figures 2.7-2.8, the comparison between the experimental composition values of the two alcohols in the distillate and at bottoms and those predicted by the model for the ternary mixture when the VLE is described by means of the values reported in the Table 2.1 are showed. In the Figure 2.7 the agreement between predicted and measured composition profiles is rather good within the first 60 minutes of operation. After this time, the model predicts an increase of the alcohol composition values, while the experimental values decrease. Note also that the alcohols appear in the residue at a time delayed of about 30 minutes with respect the experimental one, and that such a time is very close to the one corresponding to the peak of the calculated ethanol composition in the distillate. The same consideration hold for the model behavior showed in the Figure 2.8. These discrepancies are due to the incorrect calculation of the dynamics of the whole column (Baratti et al., 1998). This is due to the uncertainty of the VLE model parameters and to the evolution of temperature and composition strongly affected by the liquid dynamic hold-up.

![Figure 2.7. Comparison between ethanol and tert-butanol model prediction and off-line experimental data (Run I).](image-url)
Summarizing, the column model accounts only of the composition dynamics, while it does not describe the fast hydraulics of the column. Moreover, due to presence to close-to-azeotropic condition, and alcohols components in small amount, the estimation task in the enriching section should be considerably more difficult than in the stripping section.
Chapter 3

Nonlinear State Estimation

The state estimation for a linear system is performed adopting techniques based on a well-fixed optimal linear estimation theory. A number of sub optimal techniques have been proposed for solving the nonlinear problem. In this Chapter two techniques are presented: the EKF widely used in the chemical engineering processes; and the NGE proposed by Alvarez (1996).
3.1 Introduction

Frequently, control systems are used to model the behavior of physical, biological, or social systems

\[ \dot{x} = f(x, u) \]
\[ y = g(x, u) \]

(3.1)

where \( u \in \Omega \), a subset of \( \mathbb{R}^l \), \( x \in M \), a \( C^\infty \) connected manifold of dimension \( m \), \( y \in \mathbb{R}^n \) and \( f \) and \( g \) are \( C^\infty \) functions (Hermann and Krener, 1977).

The implementation of a control scheme requires the knowledge of the dynamic system behavior that can be obtained by means of the process model and the available on-line measured variables. There exist a number of processes in which the primary variable to be controlled is difficult to measure or it is measured with long delay in sampling and analysis process. In such cases control of the process is usually accomplished by measuring secondary variables for which sensors are more reliable, cheaper, or more readily available and installed.

In a distillation column the primary variables to be regulated are the product compositions at the bottom and in the distillate. This would involve two composition analyzers, which are often expensive and difficult to maintain. For this reason, inferential control using a soft sensor is frequently used, where the soft sensor is a dynamic algorithm (model-based) able to reconstruct the evolution of the states by means of secondary temperature measurements. To implement an inferential scheme is necessary to build the soft sensor for the unmeasured output and then use this estimated value in the control system.

The two major questions in the design of a soft sensor for a distillation column are the selection of the:

- **Estimation structure**
- **Algorithm**

Henceforth, with the term estimation structure let us indicate: (i) the number and location of the temperature sensors; (ii) the innovated/non-innovated states partition, and (iii) the innovated states per measurement; while algorithm means the dynamic data processor adopted to perform the estimation task.

It has been suggested (Foss, 1973; Stephanopoulos, 1984) that the relevant problem in chemical engineering process monitoring/control is not the development of more sophisticated algorithms, rather the establishment of a structural framework for
selecting the innovation mechanism and measured variables and linking them in an appropriately scheme. In particular, for a distillation column: what temperature should be used to estimate product compositions? How should the data assimilation mechanism be: complete or partial?

The objective of this Chapter is to focus on designing and implementing a soft sensor when the process model and on-line secondary measurements are known.

### 3.2 The Estimation Structure

Which is the best method to obtain an on-line estimation structure able to provide a good estimate of the output from the secondary measurements?

As already pointed out in the previous Section, in the distillation column field, temperatures are used as secondary measurements to control and monitor the separation process. This is the reason why the number and the temperature sensor locations have been previously studied by several authors (Boyd, 1975; Tolliver and McCune, 1980; Morari and Stephanopoulos, 1980; Yu and Luyben, 1987; Quintero-Marmol, 1991; Medjell and Skogestad 1991-1993; Baratti et al., 1995-1998; Oisiovici and Cruz, 2000-2001; Venkateswarlu and Kumar, 2006).

Temperature could be a good indicator of the corresponding concentration of the components in the product. The use of temperature to infer composition is based on the assumption of thermodynamic equilibrium between liquid and vapor on a tray of the column. At a given pressure, \( P \), there exists a unique boiling point \( \beta \) temperature corresponding

\[
\beta = \beta(c, P)
\]  

(3.2)

Assuming constant pressure and equilibrium conditions, temperature measurements will infer a binary composition. The same cannot be stated for a multicomponent system where the relationship between temperature and compositions is not uniquely defined.

The main disadvantages of using temperature as a secondary variable to monitoring distillation columns are that the selected temperature may not always well correlate with product compositions, and may not always be sufficiently sensitive to variations in product compositions.

The number and the location of the temperature sensor in a distillation column are fundamental to reconstruct the state trajectories. At this purpose the notions of system reconstructibility for linear and nonlinear systems are recalled.
3.2.1 System reconstructibility (Kwakernak and Sivan, 1972)

An important issue in choosing the estimation scheme is the number and the location of temperature sensor strictly related to the system reconstructibility.

Kwakernak and Sivan (1972) stated: It is important to know whether or not given system has the property that it is at all possible to determine from the behavior of the output what the behavior of the state is. This leads to the concept of reconstructibility of linear and nonlinear systems.

In order to discuss the reconstructibility of nonlinear systems some concepts of linear estimation are introduced.

**Definition of reconstructibility (Kwakernak and Sivan, 1972)**

Let $y(t; t_0, x_0, u)$ denotes the response of the output variable $y(t)$ of the linear differential system

$$
\dot{x}(t) = A(t)x(t) + B(t)u(t) \\
y(t) = C(t)x(t)
$$

(3.3)

to the initial state $x(t_0) = x_0$. Then the system is called completely reconstructible if for all $t_1$ there exists a $t_0$ with $-\infty < t_0 < t_1$ such that

$$
y(t; t_0, x_0, u) = y(t; t_0, x',0, u), \quad t_0 \leq t \leq t_1
$$

(3.4)

for all $u(t), t_0 \leq t \leq t_1$ implies that $x_0 = x',0$.

If a system is completely reconstructible, and the output variable is observed up to any time $t_1$, there always exists a time $t_0 < t_1$ at which the state of the system can be uniquely determined. If $x(t_0)$ is known, of course $x(t_1)$ can also be determined.

**Reconstructibility of linear systems (Kwakernak and Sivan, 1972)**

The definition of reconstructibility is due to Kalman et al. (1969). The system (3.3) is defined to be completely observable if for all $t_0$ there exists a $t_1 < \infty$ such that:

$$
y(t; t_0, x_0, u(t)) = y(t; t_0, x',0, u), \quad t_0 \leq t \leq t_1
$$

(3.5)

for all $u(t), t_0 \leq t \leq t_1$ implies that $x_0 = x',0$.

The concept of reconstructibility is complementary to the concept of observability. Observability means that it is possible to determine the state at time $t_0$ from the future output. Generally, only past output values are adopted to control and monitoring processes.
For time-invariant systems complete reconstructibility implies and is implied by complete observability.

**Observability of linear systems (Kwakernak and Sivan, 1972)**

A system is defined observable when the state can be determined uniquely from the output measurements. Observability for a linear system is a global property of the system matrices that does not depend on the specific state, input or output values.

**Theorem 3.1:** (Kwakernak and Sivan, 1972) the n-dimensional linear time invariant system, equation (3.3) is completely reconstructible if and only if the row vectors of the reconstructibility matrix

\[
Q = \begin{bmatrix}
  C \\
  CA \\
  \cdot \\
  \cdot \\
  CA^{n-1}
\end{bmatrix}
\]  

(3.6)

span the n-dimensional system.

The initial state \(x_0\) can be determined uniquely if \(Q\) is full rank. In this case, the linear system is observable or reconstructible.

Any initial state with a component in the null space of \(Q\) cannot be uniquely defined from the output measurements if \(Q\) is not full rank. This implies that if the linear system is not observable then the null space of \(Q\) is called the unobservable subspace. If the unobservable subspace does not contain unstable modes of \(A\), the system is detectable (Kwakernak and Sivan, 1972).

**Theorem 3.2:** (Kwakernak and Sivan, 1972) the unreconstructible subspace of the n-dimensional linear time-invariant system (3.3) is the null space of the reconstructibility matrix (3.6).

**Definition 3.1:** (Kwakernak and Sivan, 1972) The linear time-invariant system (3.3) is **detectable** if its unreconstructible subspace is contained in its stable subspace.

For time-varying linear system, the conditions for observability involve the state transition matrix, \(\Phi\), and the Gramian matrix, \(M\). The time-varying continuous
system is observable if only if for all \( t_1 \) there exists a \( -\infty < t_0 < t_1 \) such that the observability Gramian matrix \( M(t_0, t_1) \) is nonsingular.

The continuous time observability Gramian is a symmetric, nonnegative definite matrix computed from the state transition matrix \( \Phi(t, t_0) \).

\[
M(t, t_1) = \int_{t_0}^{t_1} \Phi^T(\tau, t)C^T(\tau)C(\tau)\Phi(\tau, t)\,d\tau. \tag{3.7}
\]

The system is completely reconstructible if the matrix \( M(t, t_1) \) is not singular.

**Observability of nonlinear systems**

The observability of a nonlinear system is locally determined about a given state or equilibrium point. Ray (1981) states that the observability of nonlinear systems is often determined by the structure of the system and is not dependent on the state in a complex manner. In this section are presented a description of the requirements for local observability of nonlinear systems (Hermann and Krener, 1977; Krener and Respondek, 1985, Krstic, 1995).

**Continuous-time nonlinear systems (Muske and Edgar, 1977)**

Let us consider a continuous time nonlinear systems described by the following nonlinear system of differential equations in which \( y(t) \in \mathbb{R}^n \) is the output, \( u(t) \in \mathbb{R}^m \) is the input, \( x(t) \in \mathbb{R}^m \) is the state of the system

\[
\begin{align*}
\dot{x}(t) &= f(x, u, t) \\
y(t) &= h(x, t)
\end{align*} \tag{3.8}
\]

Let us denote \( y(t, t_0, x_0, u(t)) \) as the output trajectory from an initial state \( x_0 \), initial time \( t_0 \), and input \( u(t) \) for the continuous-time nonlinear system.

Two states \( x_0 \) and \( z_0 \) are defined to be indistinguishable if:

\[
y(t, t_0, x_0, u(t)) = y(t, t_0, z_0, u(t)) \tag{3.9}
\]

for \( t_0 < t < T, T < \infty \), and all admissible input trajectories \( u(t) \). The nonlinear system is observable at \( x_0 \) if the set of indistinguishable from \( x_0 \) contains only \( x_0 \). (Muske and Edgar, 1977).
Local observability (Muske and Edgar; 1977)

A nonlinear system is locally weakly observable at \( x_0 \) if for every neighborhood \( \chi^n \) of \( x_0 \) there exists another neighborhood \( S^n \) contained in \( \chi^n \) such that the set of states in \( S^n \) that are indistinguishable from \( x_0 \) for all input trajectories in which \( x(t, t_0, z_0, u(t)) \in \chi^n \) and \( z_0 \in S^n \) contains only \( x_0 \).

The advantage of this definition consists on using linearized algebraic test similar to that used for linear systems.

If exists a neighborhood of \( x_0 \) and a \( p \)-tuple of integers \( (k_1, k_2, \ldots, k_p) \) (denoted as observability indices) then the nonlinear system is observable at \( x_0 \), such that:

- \( k_1 \geq k_2 \geq \ldots \geq k_p \geq 0 \quad \sum_{i=1}^{p} k_i = n \).

- The \( n \) row vectors of \( \{L_j^{i-1}(dh_i) \colon i = 1, \ldots, p; j = 1, \ldots, k_i \} \) are linearly independent.

\( L_j^{i-1}(dh_i) \) is the \( j \)-th Lie derivative (c.f., Kravaris and Kantor, 1990) of the gradient of \( h_i \) by the vector field \( f \).

An observability matrix can be constructed using as row vectors \( L_j^{i-1}(dh_i) \). The full rank of this matrix implies local weak observability.

\[
Q = \begin{bmatrix} L_j^{i-1}(dh_i) \end{bmatrix} \quad i = 1, \ldots, p; \quad j = 1, \ldots, k_i \tag{3.10}
\]

Where \( f \) and \( h \) are linear functions. The matrix is equivalent to the observability matrix for continuous-time linear systems.

It is important to note that the nonlinear observability property must be verified from the full order system.

Estimability property

Alvarez (1996) stated the following definition to determine the observability or partial observability (detectability) of a nonlinear system.

Let us consider the reconstruction, on-line, of the system state trajectories assuming that:

\[
\dot{x}(t) = f(x, u, r) \tag{3.11}
\]

\[
y(t) = h(x, r)
\]
the model parameters \((r)\) are known and that the measured signals are smooth functions of time. Using the measured signals and their time derivatives of the equation (3.11)

\[
y(t) = [y_1, \ldots, y_i^{(k_i-1)}, \ldots; y_m, \ldots, y_m^{(k_m-1)}] f(t),
\]

(3.12)

\[
k_1 + \ldots + k_m = k \leq n, \quad k_i > 0;
\]

\[
u_1 + \ldots + \nu_m = \nu \geq p, \quad \nu > 0.
\]

(3.13)

By taking successive Lie-derivatives of the output maps \(h_1, \ldots, h_m\), the following time varying, algebraic \(k\)-equation set is obtained

\[
y(t) = \phi_1 [x, u(t), r],
\]

\[
\phi_1(x, u, r) = [h_1, \ldots, L_f^{k_1-1} h_1; \ldots; h_m, \ldots, L_f^{k_m-1}]',
\]

(3.14)

\[
\phi(x, u, v, r) = [L_f^k h_1, \ldots, L_f^m h_m]'.
\]

where \(\phi_1\) and \(\phi\) are nonlinear maps.

If equation (3.14) consists of \(k \leq n\) independent equations, it establishes that the state of the plant is in the \((n-k)\)-dimensional surface, \(\Xi(t)\), which is denominated unobservable because \(k < n\).

\[
x^* = f[x^*, u(t), r], \quad x^*(t_0) = x_0, \quad x^*(t) \in \Xi(t).
\]

(3.15)

The restricted \((n-k)\) dimensional system and its motion \(x^*(t)\) will be referred to as the unobservable dynamics and motion, respectively.

**Definition 3.2:** The motion \(x(t)\) of the plant is RE (robustly-exponentially)-estimable if there are \(m\) integers (observability indices) and a map \(\phi_1(x, u, r) = [\phi_{k_1}, \ldots, \phi_{k_m}]\) such that in some neighborhood about \([x(t), u(t), r]\) (i.e. about the plant motion, the augmented input trajectory, and the parameter \(r\)):

- The map \(\phi(x, u, r) = [\phi_1, \phi_2]_{x,u,r}\) is Rx-invertible \(\phi_1 = [h_1, \ldots, L_f^{k_1-1} h_1; \ldots; h_m, L_f^{k_m} h_m]\).
- The map \(\phi(x, u, v) = [L_f^k h_1; \ldots; L_f^m h_m]\) is Lipschitz-continuous.
- The unobservable motion \(x^*(t)\) is RE-stable.
If $k = n$ the motion $x(t)$ is said to be RE-observable. If $k < n$, the motion $x(t)$ is said to be RE-detectable.

### 3.2.2 The temperature location criteria

Recalled the notion of reconstructibility of linear and nonlinear systems, an important point in the design of a soft sensor is the choice of the temperature sensor locations. The selection of sensor location to monitor batch and continuous distillation columns is still an open issue (Venkateswarlu and Avantika, 2001; Roffel, 2003; Oisiovici and Cruz, 2000-2001; Luyben, 2005; Tronci et al., 2005, Zamprogna et al., 2005; Venkateswalu and Kumar, 2006; Bian and Henson, 2006). Here, are introduced some of the criteria adopted to select the number and sensor location.

From the definitions and the theorems reported in the Section 3.2.1 it is important to note that the system reconstruction in a distillation column depends on the information content that can be obtained by means of the number and locations of measurements available. However, how many temperature sensors are necessary in order that a distillation column is observable? Where must be located the temperature sensors?

Several criteria to select the temperature sensor location have been reported in literature. Radameker et al. (1975) list several criteria, such as: linearity of response, maximum temperature, or concentration gradient, anticipatory action, minimizing composition changes inside the column. Kister (1990), in his textbook, reported the main procedures adopted to select the sensor location: (i) sensitivity, (ii) singular value decomposition (SVD), (iii) correlation with products, (iv) dynamic response.

Tolliver and McCune (1978, 1980) proposed an improved procedure, based on the sensitivity of the column temperature profile to the material balance. The column temperature profiles are evaluated at different distillate-feed ratios, while either reboiler duty or reflux flow is kept constant. The best temperature location is where temperature variations are largest and most symmetrical (Kister, 1990)

Downs and Moore (1981) introduced the use of singular value decomposition (SVD), to give guidelines to select the variables to monitor and control a distillation column. Downs and Moore used SVD to select the best tray temperatures (Luyben, 1990).

Quintero-Marmol (1991) suggested the use of singular value decomposition (SVD) to identify the most sensitive tray locations. The application of the SVD analysis for selecting the sensor locations consists of performing the singular value decomposition on a steady state gain matrix that describes the temperature
sensitivity on each tray with respect to changes in load variables. The locations that correspond to the most sensitive elements in the left singular vector are then chosen.

Oisiovici and Cruz (2001) observed that the optimal configuration obtained by applying the SVD approach to a batch column is time varying because of the dynamic behavior of the process.

The choice of temperature sensors and their location is still an open issue for any kind of estimators considered. Also this issue has been studied with a diversity of methodologies generating different point of views on the number of sensors that must be used. Undoubtedly, the singular value decomposition (SVD) is the criteria widely used to select the location of temperature measurements. In fact this procedure appears to be quite insensitive to operating conditions. Here, it is reported a briefly description of the procedure adopted in this Thesis.

**Singular value decomposition**

Many sources in the literature describe the mathematical foundations (Noble and Daniel, 1977) and the physical interpretations (Weber and Brosilow, 1972; Morari, 1983; Lau et al. 1985) of the singular value analysis. Essentially, SVD expresses a general matrix as three decomposition matrices.

SVD expresses the matrix $G$, the matrix of the plant transfer function steady state gains, as the product of three matrices: (i) the $U$ matrix, (ii) the diagonal $\Sigma$ matrix, (iii) and the $V^H$ matrix (Luyben, 1990). The matrix $G$ can be decomposed and it assume the following form (equation 3.16). More detail about the matrix decomposition can be found in Skogestad and Postlethwaite(1996):

$$G = U\Sigma V^H$$ (3.16)

Where $\Sigma (l, m)$ is matrix with $k = \min \{l, n\}$ non negative singular value, $\sigma_i$, placed in descending order along the main diagonal while the other entries are zero. The singular values are the positive square roots of the eigenvalues of $G^H G$, where $G^H$ is the complex conjugate transpose of $G$.

$$\sigma_i(G) = \sqrt{\lambda_i(G^H G)}$$ (3.17)

$U(l, l)$ is a unitary matrix of output singular vector $u_i$, $V(n, n)$ is a unitary matrix of input singular matrix vectors, $v_i$.

The ratio of the maximum, $\sigma^*$, and minimum singular value, $\sigma_*$, is the condition number of $G$. The condition number is a sensitivity measure of the system to any parameter perturbation. The minimum singular value $\sigma_*$ is a measure of the rank
deficiency or collinearity in the system and thus also a measure of invertibility (Lau et al., 1985). Therefore, $\sigma_*$ may be physically interpreted as an indicator of control effort.

### 3.2.3 The temperature sensor number

Defined where to locate the sensors another important issue is how many sensors are necessary to reconstruct the state dynamic evolution.

Yu and Luyben (1987) used a degree of freedom argument to show that a distillation column is observable as long as the number of measurements is at least $(N_c - 1)$, where $N_c$ is the mixture component number.

Quintero-Marmol (1991) stated that the nonlinear observer needed at least $N_c$ thermocouples to be effective and, for robust convergence, the number of measurements needed appears to be $N_c + 2$. Moreover, they underlined that the number of measurements required does not depend on the number of trays.

In the last years many works were devoted to the choice of the optimal sensor number and the general consensus in the control community has been that one should use only few temperatures. The complexity degree of the soft sensor design increases when more information about the process is available and this involves more susceptibility to sensor failures and measurement errors. The benefits of monitoring using multiple measurements must be carefully compared with the cost associated with building the estimator and maintaining the system. The choice frequently adopted is to consider the minimum number of measurements which is the most convenient, with respect to the soft sensor designing and tuning, and also minimizes the number of adjustable parameters.

In many instances, multiple temperature measurements are available (e.g. the continuous column considered in this work has 8 temperature measurements). Using multiple temperature measurements presents both advantages and disadvantages. The fact that generally the goodness of the estimate generally improves when measurements are added does not mean that all the measurements available should be used.

Some authors (Medjell et al., 1991-1993) have put in discussion these statements, stressing the importance of the relationship between measurements and output in order to improve the goodness of the output to be estimated.

Therefore, how many available measurements are really needed? Is it better to use only the minimum number of sensors or is it necessary to consider a possible estimate improvement increasing sensors number. Moreover, the information
CHAPTER 3. NONLINEAR STATE ESTIMATION

contained in the measurements remains constant or changes with time? For example, in the case of a high purity distillation column the main temperature drop will take place in a small region of the column. Hence, the location of the sensor in a specific region of the column could be a good or a bad choice on the basis of the evolution of the temperature front during the operating time.

3.3 Algorithms

Once defined the estimability structure (number and sensor location) it is necessary to select the algorithm. In literature there are a large variety of state estimators: static and dynamic, linear and nonlinear. However, which algorithm should be better to select? Distillation column is known to be nonlinear then is necessary to consider algorithm valid for a large range of operation. Alvarez (1999) defined four main approaches to design of nonlinear estimator:

- The Conventional Extended Kalman Filter, CEKF, whose design is simple but lacks convergence criteria and systematic tuning procedure;
- The Nonlinear Geometric Estimator (NGE), which guarantees convergence with linear output error dynamics;
- The High-Gain (HG) approach which guarantees convergence, but has a complex tuning procedure;
- The Sliding Mode (SM), which guarantees robust stability, but has an elaborated design.

3.3.1 Innovation mechanism

All the algorithms above described can be applied to completely observable plant. This means that the observability structural property, introduced in the previous Section, could be considered as a criterium to choose the algorithm. However, the NGE algorithm has the property that can be applied to observable and partially observable system. This means the possibility to reconstruct the system states by means a low order estimator and to consider the issue of the innovation mechanism, where with innovation mechanism we mean the system states that must be updated.

Recalling the equation that characterized a generic system

\[
\dot{x}(t) = f(x,u,t) + w(t) \\
y(t) = h(x,t) + v(t)
\]

(3.18)

is defined as innovation mechanism the “feedback term”:
that correct the estimate for discrepancies between the actual output data $y(t)$ and the estimated value $\hat{y}(t)$. The importance of this feedback correction is controlled by a gain matrix $G$ that depends on the error statistics of the model and of the output data. An important issue in designing of a soft sensor is to decide the state innovation set, where with innovation set we mean the set of the states to update. Hence, the choice of the innovation mechanism leads to the following questions: How should be the data assimilation mechanism? Should be complete or partial? In Chapter 4 let us show the methodology adopted to select the data assimilation mechanism in the ternary column described in Chapter 2.

Here, let us introduce the algorithm selected to design the soft sensors: the NGE and the CEKF algorithms. The CEKF can be also considered a “standard estimator” for comparison with other estimators in simulation studies.

3.3.1 The Kalman Filter

The Kalman filter is a set of mathematical equations that provides an efficient computational way to estimate the state of the process, it minimizes the mean of the squared error. The filter is powerful in estimation of past, present, and even future states (Welch and Bishop, 2001).

The Kalman filter estimates the system states at some time and then obtains feedback in the form of noisy measurements. The equations for the Kalman filter fall into two groups (Welch and Bishop, 2001):

- **Time update equations**, the current state and error covariance estimates are projected forward in time obtain a priori estimates for the next time step.
- **The measurements update**: new measurements are incorporated into the a priori estimate to obtain an improved a posteriori estimate.

Time update equations can also be considered as predictor equations, while measurement equations can be considered as corrector equations. The final estimation algorithm resembles that of predictor-corrector algorithm for solving numerical problems, as shown in the Figure 3.1 (Welch and Bishop, 2001).
CHAPTER 3. NONLINEAR STATE ESTIMATION

The Process to be estimated (Haykin, 2001)

The state is the least amount of data on the past behavior of the system that is needed to predict its future behavior (Haykin, 2001).

Let us consider the linear discrete-time dynamical system showed on the Figure 3.2 (Haykin, 2001)

\begin{equation}
    x_{k+1} = F_{k+1}x_k + w_k
\end{equation}

where $F_{k+1}$ is the transition matrix, $w_k$ is the process noise assumed to be additive, white, and Gaussian, with zero mean and with covariance matrix defined by:

\begin{equation}
    E[w_k w_k^T] = \begin{cases} Q & \text{for } n = k \\ 0 & \text{for } n \neq k \end{cases}
\end{equation}

Measurement equation (Haykin, 2001)

To estimate the unknown state, $x_k$, a set of observed data is used, denoted by the vector $y_k$. 

43
CHAPTER 3. NONLINEAR STATE ESTIMATION

\[ y_k = H_k x_k + v_k \] (3.22)

where \( y_k \) is the measure observable at time \( k \) and \( H_k \) is the measurement matrix.

\( v_k \) is the measurement noise, uncorrelated with the process noise, and it is assumed to be additive, white, Gaussian, with zero mean and with the following covariance matrix:

\[ E[ v_n' v_k' ] = \begin{cases} R & \text{for } n = k \\ 0 & \text{for } n \neq k \end{cases} \] (3.23)

The Kalman filter recursive estimation (Haykin, 2001) of the state can be summarized in the following steps shown in Fig. 3.3:

- **Initialization**
  it is possible to choose the initial estimate of the state as

\[ \hat{x}_0 = E[ x_0 ] \] (3.24)

and the initial value of the a posteriori covariance matrix as:

\[ P_0 = E[ (x_0 - E[ x_0 ])(x_0 - E[ x_0 ])^T ] \] (3.25)

This choice for the initial conditions has the advantage of yielding an unbiased estimate of the state \( x_k \).

- **State estimation propagation**

\[ \hat{x}_k^- = F_k \hat{x}_{k-1}^- \] (3.26)

- **Error covariance propagation**

\[ P_k^- = F_k^- P_{k-1}^- F_k^- + Q \] (3.27)

- **Kalman gain matrix**

\[ G_k = P_k^- H_k^T [ H_k P_k^- H_k^T + R_k ]^{-1} \] (3.28)

where \( [ H_k P_k^- H_k^T + R_k ]^{-1} \) denotes the inverse of the matrix inside the square brackets.
• **State estimate update**

\[
\hat{x}_k = \hat{x}_k^- + G(y_k - H_k \hat{x}_k^-);
\] (3.29)

• **Error covariance update**

The error covariance describes the effects of time on the covariance matrices of estimation errors.

\[
P_k = (I - G_k H_k) P_k^- .
\] (3.30)

![Figura 3.3. Kalman Filter steps (Welch and Bishop)](image)

### 3.3.2 The Extended Kalman Filter

If the model is nonlinear it is possible to extend the use of Kalman filtering by means a linearization procedure. Before to introduce the extended Kalman filter it is important to observe that the estimated variables are determined from the conditional probability density of the state. The conditional probability density of the state of a linear system (with Gaussian noise) is Gaussian and determined by the mean and the covariance. Otherwise, the conditional probability of the state of a nonlinear system is not Gaussian even when the state and measurement disturbances are Gaussian. If the system is nonlinear the determination of the conditional probability requires the use of stochastic calculus and in general it requires some approximations in order to obtain an implementable filter (Muske and Edgar, 1977). More details about probabilistic and stochastic approaches to linear and nonlinear estimation can be found in Jazwinsky (1970), Ho and Lee (1964).

The Extended Kalman filter computes a state estimate (at each sampling time) applying the Kalman filter on a linearized model of a nonlinear system. If exists a sufficiently large neighborhood where the linearized model is a good representation of the nonlinear system then this technique can be justified. Moreover, the optimal
estimate for the linearized system should be a realistic approximation if disturbances are represented by zero mean Gaussian state and measurements noise (Muske and Edgar, 1977).

Let us consider a nonlinear dynamical system described by the state space model (Haykin, 2001):

\[ x_{k+1} = f(k, x_k) + w_k \]
\[ y_k = h(k, x_k) + v_k \]  

(3.31)

where \( w_k \) and \( v_k \) are independent zero-mean white Gaussian noise processes with covariance matrices \( R \) and \( Q \). The functional \( f(k, x_k) \) denotes a nonlinear transition matrix function, while the functional \( h(k, x_k) \) denotes a nonlinear measurement matrix.

The extended Kalman Filter consists on linearizing the state-space model, equation (3.31), at each time, around the recent state estimate, \( \hat{x}_k \) or \( \hat{x}^-_k \). The approximation can be done in two steps (Haykin, 2001):

**Step 1**: The two matrices \( F_{k+1,k} \) and \( H_k \) are constructed (Haykin, 2001):

\[
F_{k+1,k} = \frac{\partial f(k, x)}{\partial x} \bigg|_{x=\hat{x}_k} \\
H_k = \frac{\partial h(k, x)}{\partial x} \bigg|_{x=\hat{x}_k} 
\]

(3.32)

\( F_{k+1,k} \) and \( H_k \) represents the partial derivative of the component of \( f(k, x) \) and \( h(k, x) \) with respect to the component \( x \).

**Step 2**: The matrices \( F(k, x_k) \) and \( H(k, x_k) \) are approximated as follows (Haykin, 2001):

\[
F(k, x_k) \approx F(x, \hat{x}_k) + F_{k+1,k}(x, \hat{x}_k), \\
H(k, x_k) \approx H(x, \hat{x}_k^-) + H_{k+1,k}(x, \hat{x}_k^-). 
\]

(3.33)

the nonlinear state equation can be approximate by:

\[
x_{k+1} \approx F_{k+1,k} x_k + w_k + d_k, \\
\bar{y}_k \approx H_k x_k + v_k, 
\]

(3.34)

where are introduced \( \bar{y}_k \) and \( d_k \):
\[ \bar{y}_k = y_k - \left[ h(x, \hat{x}_k) - H_k \hat{x}_k \right] \]
\[ d_k = f(x, \hat{x}_k) - F_{k+1,k} \hat{x}_k \]  

(3.35)

The entries in the term \( \bar{y}_k \) and \( d_k \) are all known at time \( k \).

In table 3.1 (Haykin, 2001) shows the recursions involved in computing the extended Kalman filter.

Knowing the linearized state-space model, the extended Kalman filter can be obtained applying the Kalman Filter theory.

### 3.3.3 The Tuning

The filter implementation requires the tuning if the covariance matrices \( Q \) and \( R \). The measurement noise covariance \( R \) can be obtained from the measurement data and knowledge of sensor characteristic. The determination of the model noise covariance \( Q \) is generally more difficult. The process noise covariance \( Q \) is usually selected through a trial and error procedure using computer simulation or experimental data. In many cases, a well-tuned Kalman filter can be designed by assuming a diagonal and time-invariant process noise matrix. If the required filter performance is not obtained using these procedures, an alternative to select the covariance matrices is to use a noise adaptive Kalman filter (Welch and Bishop, 2001; Oisiovici an Cruz, 2000; Oisiovi and Cruz, 2001).

In the Chapter 6 it is proposed a methodology to tune the covariance matrixes \( Q \) and \( R \) modeling the uncertainty of the equilibrium model parameters. By means the adaptive Kalman filter the covariance matrices and the state variables are simultaneously estimated.
### Table 3.1 Extended Kalman Filter (Haykin, 2001)

**State-space model**

\[
F_{k+1,k} = \frac{\partial f(x_k)}{\partial x} |_{x=x_k^{\hat{}}} \\
H_k = \frac{\partial h(x_k)}{\partial x} |_{x=x_k^{\hat{}}}
\]

**Initializations**

\[
\hat{x}_0 = E[x_0], \\
P_0 = E[(x_0 - E[x_0])(x_0 - E[x_0])^T].
\]

**State estimation propagation**

\[
\hat{x}_k = f(k, \hat{x}_{k-1});
\]

**Error covariance propagation**

\[
P_k = F_{k,k-1}P_{k-1}F_{k,k-1}^T + Q;
\]

**Kalman gain matrix**

\[
G_k = P_k^\bot H_k^T [H_k P_k^\bot H_k^T + R_k]^{-1};
\]

**State estimation update**

\[
\hat{x}_k = \hat{x}_k + G_k (y_k - f(k, \hat{x}_k));
\]

**Error covariance update**

\[
P_k = (I - G_k H_k) P_k^\bot
\]
3.3.4 The Nonlinear Geometric Estimator (NGE)

In this section are introduced: the P-NGE, PI-NGE, the estimation error dynamic, the robust convergence criteria, and the tuning procedure. The construction is based taking as reference the Alvarez works recently published (Alvarez, 1999; Hernandez and Alvarez, 2003; Lopez and Alvarez, 2004).

The P-NGE (proportional)

Let us consider the non-autonomous nonlinear plants

\[ \dot{x} = f[x,u(t),r], \quad x(t_0) = x_0, \quad x_0 \in X_0 \]

\[ x \in X, \quad r \in R \]

\[ y = h(x,r), \quad u \in U, \quad u(t) \in Y \]

with \( n \) states \( x \), \( m \) measured outputs \( y \), \( p \) known exogenous inputs \( u \); and \( n_r \) the model parameters \( r \). \( X_0 \) and \( Y \) are the sets of admissible initial states and input trajectories. The input \( u \), the state \( x \), and the parameter \( r \) take values in the sets \( U \), \( X \) and \( R \), respectively, that are compact (i.e., bounded and close) due to physical and practical considerations. Each input \( u(t) \) is a piecewise continuous function of time, and the maps \( f \) and \( h \) are sufficiently smooth in their respective domain. Thus each set \( \{x_0, u(t), r\} \) uniquely determines a state motion (Vidyasagar, 1978)

\[ x(t) = \mathcal{G}_x[t,t_0,x_0,u(t),r], \]

which in turn determines an output trajectory \( y(t) = h[x(t), r] \). The trajectory \( x(t) \) of the system describes the evolution of batch and continuous system.

Algorithm construction

Let us consider the exogenous augmented input \( Y \), equation (3.13), as the state \( x_u \)

\[ \phi = \begin{bmatrix} u_1, \ldots, u_j^{(v_1-1)}, \ldots, u_p, \ldots, u_p^{(v_p-1)} \end{bmatrix} := x_u \]

\[ x_u = \begin{bmatrix} x_{u_1}, \ldots, x_{u_1^{(v_1-1)}}, \ldots, x_{u_p}, \ldots, x_{u_p^{(v_p-1)}} \end{bmatrix} \]

of a dynamic exo-system driven by the bounded exogenous input \( u(t) \), then introduce the augmented version of the plant, where \( \Gamma_u, \Pi_u, \Delta_u \) are block diagonal matrix.
\[
\dot{x}_u = \begin{bmatrix}
    x_{u_1} & x_{u_2} & \ldots & x_{u_k} & u_{1_{\nu}} & \ldots & u_{(\nu-1)_{\nu}} & \ldots & u_{(\nu+1)_{\nu}} & \ldots & u_{(\nu)_{\nu}}
\end{bmatrix}^T
\]  
(3.40)

\[
\dot{x}_u = \Gamma_u x_u + \Pi_u v(t), \quad u = \Delta x_u,
\]
\(u(t) = [u_{1_{\nu}}, \ldots, u_{(\nu)_{\nu}}], \quad \|v(t)\| \leq \epsilon_u
\)

Let us introduce the augmented version of the plant

\[
\begin{align*}
\dot{x}_u &= \Gamma_u x_u + \Pi_u v(t), \quad u = \Delta x_u \\
\dot{x} &= f(x, \Delta x_u, r), \quad y = h(x, r).
\end{align*}
\]  
(3.41)

Recalling the observable nonlinear map \(\phi_j\), equation (3.14), let us define a non-observable map \(\phi_H(x)\) and consequently the estimability map can be defined as:

\[
\phi(x, x_u, r) = [\phi_j(x, x_u, b), \phi_H(x)]'
\]  
(3.42)

The map is continuously differentiable for \(x\) and \(u\) and is \(Rx\)-invertible (robust invertible for \(x\)) meaning that \(\phi\) and \(\phi^{-1}\) are Lipschitz-continuous.

\[
\phi^{-1}(x, x_u, r) = x(t)
\]  
(3.43)

This means that there are two sets of positive constants so that in some neighborhood of \([x(t), u(t), r]\), \(\phi\) and \(\phi^{-1}\) are bounded by:

\[
\begin{align*}
\|\phi(\hat{x}, \hat{x}_u, \beta) - \phi(x, x_u, r)\| &\leq D_x \|\hat{x} - x\| + D_\beta \|\hat{x}_u - x_u\| + D_\beta \|\beta - r\| \\
\|\phi^{-1}(\phi(\hat{x}, \hat{x}_u, \beta), \hat{x}_u, \beta) - \phi^{-1}(\phi(x, x_u, b), x_u, r)\| &\leq F_x \|\hat{x}_u - x_u\| + F_\beta \|\beta - r\|
\end{align*}
\]  
(3.44)

(3.45)

By means a coordinate change the augmented plant, equation (3.41), becomes a quasi linear system. The coordinate change \(z\) is described for the state of the exosystem and the estimability map \(\phi\) :

\[
\begin{bmatrix}
    z_u \\
    z_I \\
    z_H
\end{bmatrix} =
\begin{bmatrix}
    x_u \\
    \phi_j(x, x_u, r) \\
    \phi_H(x)
\end{bmatrix}
\]  
(3.46)

The augmented plant (3.41) in \(z\)-coordinate assumes the form:

\[
\begin{align*}
\dot{z}_u &= \Gamma_z z_u + \Pi_z v(t), \quad u = \Delta x_u; \quad \dim(z_u) = n \\
\dot{z}_I &= \Gamma_z z_I + \Pi_z \phi_j(z_I, z_H, z_u, u, b), \quad y = \Delta z_I; \quad \dim(z_I) = k \\
\dot{z}_H &= \theta(z_I, z_H, z_u, b); \quad \dim(z_H) = n - k
\end{align*}
\]  
(3.47)
CHAPTER 3. NONLINEAR STATE ESTIMATION

where \( z_u, z_I, z_H \) are the exo-observable, observable and non-observable states of the system, respectively. The observability of the matrices \((\Gamma_u, \Delta_u)\) and \((\Gamma_0, \Delta_0)\) of the augmented system, equation (3.47), and the RE-estimability of the no-observable trajectory, equation (3.15), suggests the following system as proportional-estimator of the system:

\[
\begin{align*}
\dot{z}_u &= \Gamma_u \dot{z}_u + K_u (u - \Delta_u z_u), \\
\dot{z}_I &= \Gamma_0 \dot{z}_I + \Pi_0 \phi(t, \dot{z}_I, \dot{z}_H, \nu(t), \beta) \\
\dot{z}_H &= \theta(\dot{z}_I, \dot{z}_H, \nu(t), \beta)
\end{align*}
\]

where \((\psi, \psi_0)\) are the estimates of the measured \((y, u)\). Applying the coordinate change \((\hat{x}_u, \hat{x}') = \phi^{-1}(\dot{z}, \beta)\), is obtained the nonlinear estimator:

\[
\begin{align*}
\hat{x}_u &= A_u \hat{x}_u + K_u u, \\
\hat{x}' &= f(\hat{x}, A_u \hat{x}_u, \beta) + G(\hat{x}, \hat{x}_u, \beta)[y - h(\hat{x}, \beta)], \\
\hat{y}' &= h(\hat{x})
\end{align*}
\]

where \(A_u = \Gamma_u - K_u \Delta_u\) is a stable matrix, and \(G(\hat{x}, \hat{x}_u, \beta)\) is a gain matrix given by:

\[
G(\hat{x}, \hat{x}_u, \beta) = [I_{11}, I_{21}] \phi^{-1}_x (\hat{x}, \hat{x}_u, \beta)K_0(\sigma_0)
\]

where the matrices \(I_{11}, I_{21}\) are given:

\[
I_{p[l \times n]} = \begin{bmatrix} I_{11[k \times k]} \\
I_{21[(n-k) \times k]} \\
I_{22[(n-k) \times (n-k)]} \end{bmatrix}
\]

**Estimation error dynamics**

The choice of the matrix G has been dictated by the following considerations:

- the unobservable error dynamics must not have measurements injection 
  \[
  \left[ (\partial \phi_H / \partial x)G = 0 \right],
  \]

- the observable error dynamics must be quasi linear 
  \[
  \left[ (\partial \phi_I / \partial x)G = 0 \right].
  \]

Defining the estimation error in \(z\)-coordinates \(e = [e_I, e_H, e_u, e_r]\)

\[
e_u = \dot{z}_u - z_u \text{ error of the exo-system}
\]

\[
x = I_p(x_I, x_H)', \quad (x_I, x_H)' = (I_I, I_H)x, \quad (I_I, I_H) = I_p'
\]

\^1 x = I_p(x_I, x_H)'
\[ e_t = \hat{z}_t - z_t \] observable error \hspace{1cm} (3.53)

\[ e_{\theta} = \hat{z}_{\theta} - z_{\theta} \] unobservable error \hspace{1cm} (3.54)

The error dynamic can be expressed by the following equations

\[ \dot{e}_u = A_u e_u + \Pi u \nu(t), \quad A_u = \Gamma_u - K_u x_u, \quad \mu = \Delta e_u \]
\[ \dot{e}_l = A_l e_l + \Pi_0 q_l (e_l, e_{\theta}, e_u, t), \quad A_0 = \Gamma_0 - K_0 x_{\theta}, \quad \nu = \Delta e_l \]
\[ \dot{e}_{\theta} = \omega(e_{\theta}, t) + q_{\theta} (e_l, e_{\theta}, e_u, t) \] \hspace{1cm} (3.55)

where the observability of the matrix pairs \((\Gamma_u, A_u)\) and \((\Gamma_0, A_0)\) of the augmented system, equation (3.48), guarantee the existence of gains matrix \((K_u, K_0)\) such that the matrices \((A_u, A_0)\), respectively, are stable.

Let the adjustable estimator gains have the parameterized form \((Alvarez, 1996)\)

\[ K_u(s_u) = bd \{ (s_u k_{ji}^u, \ldots, s_u k_{ji}^u) \}, \ldots, (s_u k_{i}^u, \ldots, s_u k_{ji}^u) \} \}, \quad s_u > 0 \] \hspace{1cm} (3.56)

\[ K_0(s_0) = bd \{ (s_0 k_{ji}^0, \ldots, s_0 k_{ji}^0) \}, \ldots, (s_0 k_{i}^0, \ldots, s_0 k_{ji}^0) \} \}, \quad s_0 > 0 \] \hspace{1cm} (3.57)

with entries of \(K_u\) and \(K_0\) that set stable the reference linear, non-interactive and pole assignable (LNPA) output error dynamics

\[ \mu_i^{(\nu_i)} + k_{ji}^u \mu_i^{(\nu_i-1)} + \ldots + k_{ji}^u = 0 \quad l \leq i \leq p \quad \mu_i = \psi_i^u - u_i \] \hspace{1cm} (3.58)

\[ \nu_i^{(\nu_i)} + k_{ji}^0 \nu_i^{(\nu_i-1)} + \ldots + k_{ji}^0 = 0 \quad l \leq i \leq m \quad \nu_i = \psi_i^w - y_i \] \hspace{1cm} (3.59)

the resulting parameterized matrices \(A_u(s_u)\) and \(A_0(s_0)\) meet the E-stability inequalities \((Alvarez, 1996)\):

\[ \|e^{s_u(s_u)}\| \leq a_u e^{-\lambda_u t}, \quad t \geq 0 \]
\[ \|e^{s_0(s_0)}\| \leq a_0 e^{-\lambda_0 t}, \quad t \geq 0 \] \hspace{1cm} (3.60)

where the pairs \((a_u, \lambda_u)\) and \((a_0, \lambda_0)\) are set by \(K_u\) and \(K_0\) and are independent by the parameters \(s_u\) and \(s_0\).
Convergence of the P-estimator

Theorem 3.3 (Alvarez, 1997, 1999). Let the motion \( x(t) \) of the plant, equation (3.36), be RE-estimable in the sense of the estimability property and let the reference gains, \( K_u \) and \( K_0 \), make stable the reference LNPA output error dynamics, equations (3.58)-(3.59). Then, the P-estimator:

\[
\dot{x} = A_u x_u + K_u (u - A_u \dot{x}_u), \quad \dot{u} = A_u \dot{x}_u \\
\dot{x} = f(\dot{x}, A_u \dot{x}_u, \beta) + G(\dot{x}, \dot{x}_u, \beta)[y - h(\dot{x}, \beta)], \quad \dot{y} = h(\dot{x}, \beta)
\]  
(3.61)

with the parameterized gain \( K_u \) and \( K_0 \), equations (3.56)-(3.57), and the nonlinear gain

\[
G(x, x_u, \beta) = [I_{11}, I_{21}] f(\dot{x}, \dot{x}_u, \beta)K_0(s_0),
\]

\[
[I_{11}, I_{21}] = \phi(s)^{-1}(x, x_u, \beta)
\]  
(3.62)

yields a motion \( x(t) \) that RE-converges to \( x(t) \), if the parameter \( s_0 \) is chosen sufficiently large so that:

\[
s_0 \lambda_0 > a_0 m_1 \quad \text{if } k = n
\]

\[
s_0 \lambda_0 > a_0 \left[ m_1 + c(s_0 / \lambda_1) m_1 m_1 \right] \quad \text{if } k < n
\]  
(3.63)

According to this theorem, the estimator converges if the stabilizing margin \( l_0 = s_0 \lambda_0 \) is made larger than the instability margin \( a_0 m_1 \), which accounts for a potential self-destabilization of the observable error dynamics equation (3.55).

The PI estimator

In this Section is introduced the PI (proportional integral)-NGE. The integral action is considered (in analogy with the linear control) in order to have the possibility of improving the output matching and offset reduction capabilities. The idea is to modify the P-estimator to compensate the persistent disturbance caused by the model parameter error \( e_r \). Since the exo-observable and observable error motions can be made sufficiently faster than the unobservable error motion.

The use of the integral action is justified for the following reasons: the trajectory, equation (3.55), of the exo-observable and observable errors dynamics are faster than the unobservable trajectory. The perturbation \( q_I \) of the observable error dynamic, equation (3.55), can be written in the following mode:

\[
q_I(e_I, e_{II}, e_u, e_r, t) = q_f(e_f, e_u, \nu, e_r, t) + q_e(e_{II}, e_r, t)
\]  
(3.64)
where \( q_f \) is a fast vanishing disturbance, and \( q_s \) is a slow persistent disturbances that can be regarded as an additional state \( x_q \), for which a fast estimate \( \dot{x}_q \) can be obtained from a battery of integrators (one for each output)

\[
\dot{\chi} = K_1(s_0)[y - h(\chi, \beta)] \quad \dot{\chi}_q(t_0);
\]

\[
K_1(s_0) = \text{diag}(s_0^{k_1+1}, k_1^l, ..., s_0^{k_m+1}, k_m^l).
\]  

(3.65)

The construction of the PI-NGE and its convergence are presented by the following theorem:

**Theorem 3.4** (Alvarez, 1999). Let the motion \( x(t) \) of the plant (3.36) be RE-estimable, let the gains \( K_u(s_u) \) (3.56)-(3.57) and \( G(\chi, \chi_u, \beta) \) be in the form of the P-estimator, the I-gain \( K_I(s_0) \) be given by 3.65 and let the entries \( K_u, K_0, K_I \) set stable the reference LNPA output error dynamics (3.58)-(3.59), respectively. Then the PI-estimator expressed by the following equations:

\[
\dot{\chi}_u = \Theta_1(\chi, \chi_u, t_0) + K_I(\dot{\psi}_u - \chi)
\]

\[
\chi = f(\chi, \chi_u, \beta) + G(\chi, \chi_u, \beta)[y - h(\chi, \beta)] + G(\chi, \chi_u, \beta) \int_{t_0}^{t} \left[ \dot{y} - h(\chi, \beta) \right] dt
\]

where \( G \) is the proportional gain and \( H \) is the integral gain defined as

\[
H(\dot{x}, \dot{x}_u, \beta) = [I_{11}, I_{21}, I_{21}^{-1}] \phi^{-1}_{x1}(\dot{x}, \dot{x}_u, \beta) \Pi_0
\]  

(3.67)

**Convergence of the PI-estimator**

The PI-estimator (3.67) yield a RE-convergent estimate \( \dot{x}(t) \) of \( x(t) \) if the parameter \( s_0 \) is chosen sufficiently large so that the following inequality is met:

\[
s_0 > a_0 m_1 \quad \text{if } k = n
\]

\[
s_0 > a_0 [m_1 + c_1^{s}(a_x / \lambda_n, m_{II} n_1)] \quad \text{if } k < n
\]  

(3.68)

(3.69)

where \( \{m_1, m_{II}, n_1, a_x / \lambda_n, c_1^{s}\} \) are defined in Theorem 3.3.
**NGE-Tuning**

The nonlinear error in the observable error dynamics should be sufficiently small, in order that the convergence condition would be easily met, and the estimator divergence would not be a problem and the estimator gain matrix $K_0$ should be tuned with the noise oriented pole placement approach.

The $i$-th reference LNPA output error equation has the following expression:

$$v_i^{(k_i)} + k_{ii}v_i^{(k_i-1)} + \ldots + k_{i1}v_i = 0 \quad n_i = k_i$$  \hspace{1cm} (3.70)

and require its characteristic polynomial

$$(\gamma_i)^n + k_i^1(\gamma_i)^{n_i-1} + \ldots + k_i^{n_i} = 0$$  \hspace{1cm} (3.71)

to have $n_i$ prescribed reference poles $\gamma_i^1, \ldots, \gamma_i^{n_i}$, according to the following robustness-oriented criteria: (i) the condition number must be one, in order to attenuate the effect of the disturbance of the error dynamics, and (ii) each pole must have a sufficiently large damping factor to prevent excessively oscillatory behavior. Thus if $n_i$ is an even number, the reference LNPA output error dynamics must have $n_i/2$ repeated pairs of complex poles with reference characteristic frequency $\omega_i^0$ and a sufficiently large damping factor.
Soft Sensor Design
Introduction

In Chapter 3 was underlined that Soft Sensor design involves structural and algorithmic decisions. In the next two Chapters is to develop a methodology to select a robust the estimation structure that jointly combined with an algorithm must be able to reconstruct the state trajectories.

In the distillation column estimation field the structural property, i.e. the choice of temperature sensors number and their location, is still an open issue that has been studied with a variety of methodologies generating different point of views on the number of sensor that must be used. Indeed, while some studies state that only a few (two or three) temperature sensors should be used in order to prevent ill-conditioning by an over-parameterized problem (c.f., Yu and Luyben, 1987; Quintero et al., 1991; Baratti et al., 1998), other studies questioned this claim (c.f., Mejdell et al., 1993) pointing out the importance is not in the sensor number but in the relationship between measurements and output that must be estimated. The choice of the estimation structure has a great importance in the soft sensor design, however the problem related to the choice of the algorithm to adopt must not be neglected.

Generally, the Conventional Extend Kalman filter (CEKF) is by far the most widely used technique to estimate concentrations on the basis of temperature measurements. On the other hand, the rapid growth of the CEKF dimensionality with the number of components and stages motivates the development of a unified framework to jointly address the structure and algorithm estimation designs, with emphasis on the attainment of schemes with measurement structure selection criterion, suitable data assimilation, and as much as possible simplified algorithms.

In Chapter 4 is illustrated the methodology adopted to develop an estimation structure able to monitor the system states with a reduced dimension of the innovated (estimate) state vector and taking the maximum advantages from the information obtained by the selected temperature.

In particular, the approach is pursued with emphasis on: (i) regarding the estimation structure: (i.e., the innovated/non-innovated partition, number of sensors and their location) as a design degree of freedom, (ii) finding out to which extent the estimator functioning depends or not on the choice of estimation algorithm selected, (iii) developing an adaptive estimation scheme to perform the data assimilation task in a more efficient way.
In Chapter 5, the resulting data assimilation structure is implemented with nonlinear geometric (NGE) and EKF algorithms. In both cases, the estimation task is performed with an algorithm that is considerably simpler (measured by the number of dynamic equations and the simplicity of the tuning scheme) than the one of a conventional CEKF with complete data innovation injection scheme.
Part I

Estimation Structure Design
Chapter 4

The Data Assimilation Mechanism

In this Chapter, a structure-oriented approach to draw the data assimilation scheme is developed. The methodology presented includes the choice of sensor numbers and locations as well as the set of innovated states where the measurement information is directly injected.
4.1 The Estimation Problem

As introduced in the previous Chapters the main issues in a column estimator design are the choice of: (i) structure (i.e., number and location of the temperature sensor, and innovated state set); and (ii) algorithm. To design any algorithm it is necessary to test the structural observability (or detectability) property illustrated in Chapter 3. Indeed, the knowledge of the system observability is important to design an algorithm able to reconstruct the states dynamic by means of the injection of measured output. Theoretically speaking, a distillation column is observable, at the steady state condition, if the number of temperature measurements is equal to \((N_c - 1)\), where \(N_c\) is the mixture component number (Yu and Luyben, 1987).

Sometimes, although the observability property is met, the construction of a full order soft sensor for systems characterized by a large number of balance equations could become very complex from a computational point of view and the complexity degree can increase because of the algorithm chosen. For example, the soft sensor design for a system of \(N_c\) components, observable, characterized by a state vector of large dimension and with \((N_c - 1)\) temperature sensor, could be very cumbersome considering that: (i) the dimension of the gain matrix depends on the dimension of the state vector and on the number of the temperature sensor; (ii) and changing the column operating condition could be possible that during the column transient the observability property may poorly met at least in some location.

Theoretically to reconstruct the states dynamic evolutions in a ternary system, as considered in this Thesis, are necessary two temperature measurements. However, recalling the thermodynamic considerations stated in Chapter 2, during the transient the column moves closely a region where the relationship temperature-composition is weaker, consequently the observability property is weakly met compromising the composition estimation quality. From the geometric estimation approach (Alvarez et al., 1999; Alvarez, 2000), we know that if a nonlinear system is nominally completely observable in an proper (suitable) sense (Lopez et al., 2004), it may be convenient to perform the measurement innovation only in a subset of states in order to execute the data assimilation task in a simpler and more robust way. The underlying notion of robust detectability corresponds to a coordinate-dependent form of an instantaneous observability-based definition of detectability (Fernandez, 2007). In other words, the states that are nominally but not robustly observable should be transferred to a vector of noninnovated states, and this enables a design degree of freedom to look for a suitable compromise between data assimilation and (measurement and model) error propagation.
To solve the estimation problem and to choose the estimability structure better suited for the ternary system considered it is necessary recall the balance equations for a component on the generic i-th tray:

\[ c_{\cdot i} = \left[ (R + F) \Delta c_{\cdot i} - V \Delta y_{\cdot i} f_k,i \right] / \eta_i = f_k,i \]  
(4.1a)

\[ y_i = T_i = \beta(c_i), \quad i = s_{n_1}, \ldots, s_{n_T} \]  
(4.1b)

\[ u(t) = \{V, R, F, c_{i-1}, c_{i+1}, c_{\cdot i} \} \]  
(4.1c)

Our estimation problem consists in finding the estimation structure that yields the best compromise between reconstruction rate and tolerance to model errors. In particular we are interested in: (i) identifying the interplay between structure and error propagation, (ii) defining a priori (before implementation) estimability measures to discriminate structures and (iii) putting the adjustable approach in perspective with the conventional fixed structure CEKF algorithm.

### 4.2 Adjustable-Structure Assessment

The choice of the estimation scheme better suited for the ternary system considered is obtained by means of nonlinear detectability measures. Here, the detectability measures are defined to assess the reconstruction rate and the error propagation via the innovated and noninnovated estimation dynamics. The application of these measures in the light of physical considerations yields a reduced set of candidate structures.

#### 4.2.1 Detectability structures

Here, let us recall the definition of nonlinear detectability and estimability presented by Lopez and Alvarez (2004). The observability index set, \( i \), (Definition 2, Chapter 3) determines the set \( \Omega \) of the admissible estimator structures \( \Sigma \):

\[ \Omega(i) = E(i) \times X_p(i), \]  
(4.2a)

\[ E(i) = \left\{ k = (k_1, \ldots, k_m) \leq n \right\}, \]  
(4.2b)

\[ X_p = \left\{ x_p \in \varepsilon(x_o) \right\}, \]  
(4.2c)

\[ \Sigma = (k, x_p) \in \Omega(i) \]  
(4.2d)

where \( k \) is the estimator degree vector, \( x_p \) is the vector of states transferred from the observable state \( x_o \) to the unobservable one \( x_{\cdot p} \). \( E(i) \) and \( X_p(i) \) are respectively the set of admissible estimation degree vectors and that of transferable states. The choice of
CHAPTER 4. ESTIMATION STRUCTURE DESIGN

a structure $\Sigma$ determines the innovated (or noninnovated) state $x_I$ (or $x_{II}$), the noninnovated dynamics (4.3e) and their unique solution or noninnovated motion $x_{II}^*$.

\[
x_I = (x_I, \ldots, x_k), \quad x_{II} = (x_{II}, \ldots, x_k),
\]

\[
x_p = (x_{k+1}, \ldots, x_K), \quad \text{dim}(x_I, x_{II}) = (k, n-k),
\]

\[O(x_I, x_p, u, k) = \partial_{x_I} \Phi(x_I, x_p, u, k),\]

\[\phi = (\phi_1, \ldots, \phi_k), \quad \phi_k = (h_1, \ldots, L_{x_I}^{-1}h_1)
\]

\[
x_{II}^* = f_{II}^{-1}(x_{II}, x_{II}^*, u, k), x_{II}^*, x_{II}^* = f_{II}^{-1}[x_{II}, x_I(t), u(t), k]
\]

\[
x_{II}^*(t) = \Phi_{II}^*[t, t_0, x_{II0}, x_I, u, k], \quad x_{II}^*(t_0) = x_{II0}
\]

\[
x_I(t) = [y_I, \ldots, y_{k-1}, y_2, \ldots, y_{k-1}^2, \ldots, y_3, \ldots, y_3^k, y_4]
\]

where $\partial_{x_I} \Phi = \partial \Phi / \partial x_I$, and $L_f(\alpha)$ is the $i$-th directional derivative of the scalar field $\alpha$ along the vector $f$:

\[
L_f^i \alpha = L_f(L_f^i \alpha), \quad L_f \alpha = \partial \alpha + \partial \alpha, \quad L_f^0 \alpha = \alpha
\]

According to the estimation approach (Alvarez and Lopez, 1999; Alvarez, 2000; Hernandez and Alvarez, 2003), the column motion $x(t)$ is robustly estimable with structure $\Sigma$ if:

I. $\text{rank } O[x_I(t), x_I(t), u(t), k] = k \quad \forall t$

II. The noninnovated motion $x_{II}^*(t)$ is stable.

4.2.2 Detectability measures

In order to select the optimal estimation structure Lopez and Alvarez (2004) proposed the following guidelines:

- To minimize the error propagation via the fast innovated dynamics, the structure $\Sigma$ should be chosen to minimize the condition number, $c_{x_I}$, of the innovation matrix $O$:

\[
c_{x_I}(t) = \text{cn}[O[x_I(t), x_I(t), u(t), k]]
\]

- To minimize the error propagation via the slow noninnovated dynamics, $\Sigma$ should be chosen to minimize the stiffness measure $c_{x_{II}}$ of the map $f_{II}^*$.
\[ c_{sII}^{f}(t) = cn[J(t)], \quad J(t) = [\hat{\mathbf{X}}_{sII} f_{sII}^{*}][x_{II}(t), x_{II}(t), u(t), k] \]  
(4.5b)

- To maximize the reconstruction rate via the estimator convergence at maximum speed, \( \Sigma \) should be chosen to maximize the time varying stability margin \( I_{II} \) drawn from the Lyapunov-based sufficient condition for the exponential stability of the noninnovated motion \( x_{II}^{*}(t) \):

\[
l_{II}(t) = (1/2)\min\{-\lambda_{1}(t),...,\lambda_{n-k}(t)\} > 0
\]
(4.5c)

\[
\lambda_{i}(t) = (1/2)ev_{i}[J']', \quad J'(t) = [J_{sII}(t) + J'(t)]
\]
(4.5d)

where \( cn \) (or \( ev_{i} \)) is the condition number (or i-th real eigenvalue) of the matrices \( O, J \) and \( J' \).

These guidelines suggest measuring the estimability property of the column by means of the three-measure function set:

\[
M_{k}(t) = \left\{ c_{sII}^{f}, c_{sII}^{e}, I_{II}(t) \right\}
\]
(4.5e)

where \( c_{sII}^{f}, c_{sII}^{e}, I_{II} \) are the condition number with respect the innovation and noninnovated matrix and the stability margin. The condition number, \( c \), or the sensitivity measure \( S \) (the inverse of the minimum singular value of the matrix) provides information on the robustness of the estimation structure, while the stability margin provides information related to the stability of the noninnovated state trajectory:

\[
c = \frac{\sigma'}{\sigma}, \quad S = \frac{1}{\sigma}
\]
(4.5f)

A large condition number means weak robustness and consequently a weak soft sensor convergence and small tolerance to input and model parameters errors. For this reason the detectability measures are adopted as tools that allow individuating the estimation structure with the best compromise between reconstruction rate and robustness (tolerance to model errors).

### 4.3 Structural Assessment

From the literature is known that a ternary column is completely locally observable about a steady state, (Yu and Luyben, 1987; Quintero-Marmol et al., 1991) with at least two temperature sensors. This assessment should be revised in the light of a nonlinear instantaneous observability framework (Alvarez and Lopez, 1999,
Alvarez, 2000, Lopez and Alvarez, 2004), and a comment on the consideration of a single-sensor must be done. In a way that is analogous to the design of robust controllers via backstepping (Krstic et al., 1995, Lopez and Alvarez, 2004), in the geometric estimation approach one gives up the (possibly ill-conditioned) complete (nominal) observability structure, in order to favor robustness, and diminish observability requirements, at the cost of a more sluggish reconstruction rate. From this viewpoint, it makes sense to consider a multicomponent distillation column (possibly with ill conditioned complete observability property) with a robustly detectable single-sensor structure. In the next Section general guidelines to individuate a robust estimation structure are presented.

4.3.1 The single tray detectability analysis

The objective of this Section is to propose a methodology, based on the geometric estimation approach, able to find the estimation scheme better suited for a generic multicomponent column. For this purpose, let us consider a system constituted by a generic mixture of \((N_c-1)\) components and one temperature sensor located in the generic i-th tray.

Generally, an estimation structure is indicated with the following notation (Lopez and Alvarez, 2004)

\[ \Sigma = (k, x_I, x_{II}) \in \Omega(i) \]  

(4.6)

Recalling the detectability measures introduced in previous Section, equations (4.5), and regarding the state set \(\{c_{1,i-1}, c_{2,i-1}, ..., c_{N_c-1,i-1}, c_{1,i+1}, c_{2,i+1}, ..., c_{N_c-1,i+1}\}\) associated with the generic i-th stage, let us consider the admissible candidate structures (\(\Sigma_i\)):

\[ \Sigma_1 = \{k = N_c-1, x_I = N_c-1, x_{II} = \emptyset\} \in \Omega(i), i = \{k_1\} \]  

(4.7a)

\[ \Sigma_2 = \{k = N_c-2, x_I = c_{1,i}, x_{II} = c_{P,i}\} \in \Omega(i), i = \{k_1\}, P \neq I \]  

(4.7b)

\[ \Sigma_3 = \{k = (N_c-2, N_c-2), x_I = (c_{1,i}, c_{2,i}, ..., c_{N_c-1,i})\} \in \Omega(i), i = \{k_1\} \]  

(4.7c)

and their corresponding singularity measures

\[ S_1 = \frac{1}{\text{msv} \Omega(c_{1,i}, c_{1,i})} \]  

(4.8a)

\[ S_2 = \frac{1}{\beta(c_{1,i}, c_{2,i}, c_{N_c-1,i})} \]  

(4.8b)

\[ S_3 = \frac{1}{\text{msv} \Omega(c_{1,i}, c_{2,i})} \]  

(4.8c)
where \( msV \) denotes the “minimum singular value”, and \( O \) is the innovation matrix. The subscript \( j \) associated to the generic estimability structure \((\Sigma_j)\) denotes the structure corresponding to the number of temperature sensors (number and location) and to the innovated/noninnovated state set chosen (corresponding to the estimability degree selected).

Henceforth, in this Thesis let us adopt the following scheme to indicate the estimation structure:

\[
\Sigma_k = \Sigma_k (\pi_{ij}, x_I, x_{II}), \quad i = 1, \ldots, m \quad j = 1, \ldots, n
\]  

(4.9)

where \( \pi_{ij} \) represents the number of temperature sensors (subscript \( i \)) and the sensor location (subscript \( j \)), \( x_I \) (or \( x_{II} \)) is the innovated (or not-innovated) state. It is important to note that the estimability degree \( k \) is strictly related with the sensor number and location (\( \pi_{ij} \)).

The sensitivity measures provide information about the ill conditioning of the innovated dynamics (equations 4.5a). However, in order to select a robust estimation structure it is important to analyze the behavior and the stability of the noninnovated states; this is possible recalling the equation (4.5c). Indeed, additional information on the choice of the estimation structure can be found analyzing the behavior of the innovated states and the dynamic parameter \((\widetilde{\lambda}_{II,i})\), which stands for exponential stability margin of the noninnovated states and provides the response speed to reconstruct the states dynamic via estimator. In the following equations it is possible to observe that the dynamic parameter, \( \widetilde{\lambda}_{II,i} \), depends on two contributes: (i) the natural dynamic of the noninnovated state \((\lambda_n,i)\), and (ii) the innovated dynamic \((\lambda_{I,i})\),

\[
\lambda_{II,i} = - (\lambda_n,i + \lambda_{I,i}), \quad \lambda_n,i = \frac{\partial f_{P,i}}{\partial c_{P,i}} \\
\lambda_{I,i} = (\frac{\partial f_{P,i}}{\partial c_{I,i}})(\frac{\partial c_{I,i}}{\partial c_{P,i}}), \quad \frac{\partial c_{I,i}}{\partial c_{P,i}} = \frac{\partial \beta_{c,I}}{\partial \beta_{c,P}}
\]

(4.10a,b,c,d)

As can be seen from the equations (4.10) the influence of the innovated dynamic on the noninnovated depends on the ratio of the bubble point temperature derivatives. The influence of this ratio on the noninnovated dynamic suggests: (i) the composition state that should not be innovated when there is a strong interaction between the innovated (I) and noninnovated dynamics (P); and (ii) the estimation structure that should be used when the innovated dynamic has weak influences on the noninnovated dynamics.
Defined the sensitivity measures, the next step will be to apply these measures as tools that jointly with the thermodynamic considerations presented in Chapter 2 permit to individuate the estimation structure characterized by the best compromise between reconstruction rate and robustness.

4.3.2 Passivity concepts

Before starting the sensitivity analysis let us recall the notion of dissipative and passivity system, more additional details can be found in Willems (1972) and Byrnes (1990-1991).

Passivity systems (Willems, 1972) appear in several areas of control theory. A system is defined dissipative if it is unable to produce a specified quantity, such as energy. Here, let us introduce the fundamental theorems concerning passivity and let us adopt them as tools to solve the estimation problem.

Let us consider a nonlinear system described by the equation of the form (Willems, 1972):

\[
\begin{align*}
\dot{x} &= f(x) + g(x)u \\
y &= h(x)
\end{align*}
\]

With state space \( X \in \mathbb{R}^n \), set of inputs values \( U \in \mathbb{R}^m \) and set of output values \( Y \in \mathbb{R}^m \). The set \( Y \) of admissible inputs consists of all \( U \)-valued piecewise continuous functions defined on \( \mathbb{R} \). \( f \) and the \( m \) columns of \( g \) (i.e. \( C^\infty \)) are smooth vector fields and \( h \) is a smooth mapping. Let us suppose that the vector field \( f \) has at least one equilibrium, thus it possible to assume \( f(0) = 0 \) and \( h(0) = 0 \).

**Dissipative system,**

Let \( w \) be a real-valued function defined on \( U \times Y \), called the supply rate.

We assume that for any \( u \in Y \) and for any \( x^o \in X \), the output \( y(t) = h(\Phi(t, x^o, u)) \) of (4.11) is such that \( w(s) = (w(u(s), y(s))) \) satisfies

\[
\int_0^t |w(s)| ds < \infty \quad \text{for all } t \geq 0.
\]

**Definition 4.1:** A system of the form (4.11) with supply rate \( w \) is said to be dissipative if there exists a \( C^0 \) nonnegative function \( V: X \to \mathbb{R} \), called the storage function, such that for all \( u \in Y \), \( t \geq 0 \)

\[
V(x) - V(x^o) \leq \int_0^t w(s) ds
\]

where \( x = \Phi(t, x^o, u) \).
Passive system

A system $\Sigma$ of the form (4.11) is said to be passive if there exists a $C^\infty$ nonnegative function $V: \mathbb{R}^n \rightarrow \mathbb{R}$, called the storage function, which satisfies $V(0) = 0$, such that for all $u \in \mathbb{R}^n$, $t \geq 0$

$$V(x) - V(x^0) \leq \int_0^t \dot{V}(s) ds.$$  \hspace{1cm} (4.14)

it follows that passive system having a positive definite storage function $V$ are Lyapunov stable. (Byrnes et al., 1991).

In this Thesis the concept of passive system is extend to passive structure in the sense that their estimation degrees are equal to one and the noninnovated dynamic is stable. These concepts are expressed by the following theorems:

**Theorem 4.1** (Byrnes et al., 1990): Suppose $\Sigma$ is a passive system with a $C^2$ storage function $V$ which is positive definite. Suppose that $x = 0$ is a regular point for $\Sigma$, then $L_y h(0)$ is non-singular and $\Sigma$ has relative degree at $x = 0$.

**Theorem 4.2** (Byrnes et al., 1990): Suppose $\Sigma$ is a passive system with a $C^2$ storage function $V$ which is positive definite. Suppose that $x = 0$ is a regular point for $\Sigma$, then the zero dynamic of $\Sigma$ locally exists at $x = 0$ and $\Sigma$ is weakly minimum phase.

The Theorems 4.1 and 4.2 show that any passive system with a positive storage function has a relative degree $\{1, \ldots, 1\}$ at $x = 0$ and is weakly minimum phase. Here, let us remark that a system called at minimum phase is a system whose zero dynamics are asymptotically stable.

Further motivations for the employment of a passive estimation structure in a combined estimation-control passive design can bee seen elsewhere (Byrnes et al., 1991; Krstic et al., 1995; Alvarez et al., 2004; Gonzalez and Alvarez, 2005; Alvarez et al., 2005).
4.4 The Ternary Column

In this Section, the detectability measures previously introduced are applied to test the estimation structure better suited for the ternary column presented in Chapter 2.

4.4.1 The coupled structure

Ethanol and tert-butanol as innovated states

Supposing to know the states \{c_{E,i-1}, c_{T,i-1}, c_{E,i+1}, c_{T,i+1}\} and locating the single temperature sensor in a generic i-th stage, the estimation structure assumes the following expression:

\[ \Sigma_i = \{ \pi_{ij}, x_i = (c_{E,i}, c_{T,i})' \} \]  

(4.15a)

The observability matrix \( O \), corresponding to the estimability structure above mentioned, is made up by two row vectors.

\[
O = \left[ \begin{array}{c} \beta_{c_{E}} (c_{E,i}) \\ \beta_{c_{T}} (c_{E,i}) \\ \partial \phi (c_{E,i}, c_{E,i-1, i}, c_{E,i+1, i}, u) / \partial c_{E,i} \\ \partial \phi (c_{E,i}, c_{E,i-1, i}, c_{E,i+1, i}, u) / \partial c_{T,i} 
\end{array} \right]
\]

(4.15b)

The components of the first row vector (4.15b) express the influence of thermodynamic on the system while the components of the second row vector (4.15b) express the influence of the thermodynamics combined with the operating condition. This means that the system observability matrix does not depend only from the system thermodynamic but also from the system operating conditions. Moreover, the interactions between the matrix entries can strongly influence the matrix conditioning. If the observability matrix is near to the singularity condition the error propagation grows. For this reason is suggested to study the system sensitivity. The sensitivity is defined as the inverse of the minimum singular value of the observability matrix and it gives a measure of the error propagation along the column.
In the case study considered the sensitivity measure is expressed by the following equation:

\[ S_{E-T,i} = \frac{1}{m} \text{sv}_O (c_i, c_{i-1}, c_{i+1}, u) \]  

(4.15e)

where \( S_{E-T,i} \) is the error propagation measure for a sensor located in the \( i \)-th stage.

The sensitivity measure has been calculated for all the experimental runs reported in Chapter 2 (Table 2.4). However, the sensitivity analysis is quite insensitive to operating conditions. This is the reason for which is showed only the sensitivity measure referred to one of the experiments. The Figure 4.1 shows the sensitivity measure referred to the Run I. The Figure shows the evolution of the sensitivity measure during three times: initial, intermediate, and final time. As it can be seen, when the process separation is started, the sensitivity assumes small values along the column, but during the process evolution the sensitivity increase in the stripping and in the enriching section. The column region with lowest values of sensitivity measure is the column bottom.

![Figure 4.1. Singularity measure dependency when ethanol and tert-butanol are innovated state, at three different times.](image)

This suggests that if a soft sensor with the estimation structure \( \Sigma_i \) must be designed could be better to locate the sensor at the stage 0. In addition, it is important to remark that in the remaining column stage column the two-innovated state error propagation measure is considerably larger, meaning that, the information injected by the measured temperature is not sufficient to reconstruct the states dynamic. This is due to: (i) the combination of interactions in \( O \), (ii) the presence of first and second order partial derivatives of the equilibrium (\( \nu_E \) and \( \nu_T \)) and bubble point (\( \beta \)) nonlinear functions in \( O \), (iii) the stages with close-to-azeotropic compositions, and dependency of \( O \) on neighbor tray compositions. These results lead us to disregard...
CHAPTER 4. ESTIMATION STRUCTURE DESIGN

nonpassive estimation structures or if an estimator with this structure must be
designed the temperature sensor should be located in the column bottom.

4.4.1 Passive structures

Let us continue the structural assessment by considering single-sensor robustness-
oriented passive structures with one innovated state.

Let us denote the two estimation structures according with the following notation:

\[ \Sigma_2 = (\Sigma_{2a}, \Sigma_{2b}) \]  (4.16a)

\[ \Sigma_{2a} = \{ \pi_{ij}, x_I = (c, T_i^2), x_{II} = (c, T_i^2) \} \]  (4.16b)

\[ \Sigma_{2b} = \{ \pi_{ij}, x_I = (c, T_i^2), x_{II} = (c, E_i^2) \} \]  (4.16c)

where with \( \Sigma_{2a} \) is denoted the structure with ethanol as the innovated state, and with
\( \Sigma_{2b} \) the structure where tert-butanol is the innovated state.

**Ethanol as innovated state**

Let us consider the estimation structure (4.16b) characterized by a single sensor
single-innovated state estimation scheme. As in the previous case one temperature
sensor located in the generic i-th stage is considered while only ethanol is considered
as innovated state. The observability matrix is characterized by only one entry and
consequently the minimum singular values coincide with the temperature derivative
of the bubble point relationship.

\[ O(c, T_i^2, c, T_i^2) = \beta_{cE}(c, T_i^2, c, T_i^2) \]  (4.17a)

\[ S_E^i = 1/|\beta_{cE}(c, T_i^2, c, T_i^2)| \]  (4.17b)

The Figure 4.2 shows the sensitivity measure dependency on the spatial (stage)
location, for three different times: initial, intermediate, and final time.

As it is possible to observe, during the process evolution, the sensitivity measure
increases with time especially in the enriching section. As expected, recalling the
weak relationship between temperature-and composition for the system ethanol-
water, this is due to the presence, in the enriching section, of a region where small
temperature changes imply large composition variation. Since, the stripping section
is the region with the lowest values of the sensitivity measure and therefore
characterized by a good data assimilation capability. This suggests locating the
temperature sensor in this region. However, the difference between the sensitivity
values reached in the enriching and in the stripping section is not so marked
especially close the column middle. This implies that possible location for the
temperature sensor could be both the enriching and the stripping section.
Recalling the detectability measures introduced in the previous Section, after the analysis of the sensitivity related to the innovated states is necessary to analyze the sensitivity connected to the noninnovated states and the stability margin parameter.

Analyzing the behavior of the noninnovated states and the dynamic parameter, $\lambda_{~II,i}$, it is possible to obtain useful information about the robustness of the estimation structure selected.

$$\lambda_{~II,i} = - \left( \lambda_{n,i} + \lambda_{~I,i} \right)$$

where

$$\lambda_{n,i} = \frac{\partial f_T,i}{\partial c_T,i}$$

and

$$\lambda_{~I,i} = \left( \frac{\partial f_T,i}{\partial c_E,i} \right) \left( \frac{\partial c_E,i}{\partial c_T,i} \right)$$

As can be seen from the equations (4.18c,d) the influence of the innovated dynamic on the noninnovated one depends on the ratio of the bubble point temperature derivatives. Therefore, the two dynamics, natural and noninnovated, is showed in the Figures (4.3a,b) for three different times. It is important to observe that the innovated dynamic weakly influences the noninnovated dynamic. This suggests that the application of this single-estimation structure could be a good choice in the stripping and in the enriching section.
Chapter 4. Estimation Structure Design

Figure 4.3.a,b. (a) Natural dynamic of the not-innovated state, (b) Influence of the innovated state on the not-innovated dynamic.

The stability margin parameter provides information about the stability of the noninnovated trajectory and it is shown in the Figure 4.4. The case study considered confirms that the estimation scheme with ethanol as innovated state could be a possible candidate estimation structure.

Figure 4.4. Stability margin parameter at three different times.

Tert-butanol as innovated state

The same analysis can be performed for the tert-butanol component. The single-sensor single-state innovation scheme (4.16c) is characterized by an observability matrix have the following scheme

\[ O(c, c^E, c^T) = \beta(c^E, c^T) \] (4.19a)

As in the previous case the observability matrix is characterized by only one entry. Consequently the minimum and maximum singular values coincide with the temperature derivative of the bubble point relationship with respect to tert-butanol composition.
\[ S^T_{i} = 1/|\beta_{cT}(c_{E}^{\sim E}, c_{E}^{\sim T})| \]  

(4.19b)

The Figure 4.5 shows the sensitivity measure referred to the estimation scheme (4.16c) at three different times: initial, intermediate, and final time. It is possible to observe that the information obtained by the injection mechanism of the measured temperature is not working in the enriching sections. The reasons must be seen in the weak relationship between temperature and composition in the stage [20-30] of the enriching section. According to the results depicted in Figure 4.5, the tert-butanol should definitely not be chosen as the innovated state when the measurement is located about tray 23, and the same structure should not chosen for the stage interval 20-30.

Moreover, the analysis of Figures (4.6a,b), according with the equations (4.20) suggests not innovate this state in the enriching section considering that the singularity measure assumes large values and exist a strong interaction between the innovated and the noninnovated dynamic.

\[ \lambda_{n,i} = \frac{\partial f_{E,i}}{\partial c_{E,i}} \]

(4.20a)

\[ \lambda_{I,i} = (\frac{\partial f_{E,i}}{\partial c_{E,i}})(\frac{\partial c_{E,i}}{\partial c_{E,i}}), \quad \frac{\partial c_{E,i}}{\partial c_{E,i}} = \frac{\partial \beta_{cE}}{\partial \beta_{cT}} \]

(4.20b)

As expected the stability margin parameter shows a marked instable behavior, see Figure 4.7, in the enriching section between the stages [20-30]. Finally, we can conclude observing that the implementation of single-model estimation could be a good choice in the stripping section that is characterized by a good data assimilation.
capability, while the same consideration cannot be done for the enriching section with a weak relationship between temperature and composition.

Figure 4.6.a,b, (a) Natural dynamic of the noninnovated state, (b) Influence of the innovated state on the noninnovated dynamic

Figure 4.7. Stability margin parameter.

**Remarks**

Comparing Figures 4.2 and 4.5, when the measurement is located in the stage interval 0-16 either the ethanol or the tert-butanol can be chosen as innovated state, and in both cases sensors location in the interval 20-30 should be avoided, especially for the case of tert-butanol as innovated state. Physically speaking this means that: (i) in the stage interval 0-12 the relationship temperature-composition is sufficiently strong for estimation purposes, with ethanol or tert-butanol as innovated state, and (ii) in the stage interval 20-30 the estimation task via measurement injection is more difficult, could be pursued with the ethanol as innovated state but not with the tert-butanol as innovated state, because the presence of ethanol (or tert-butanol) is mildly (or imperceptible) reflected in the temperature measurement. As showed in Figures 4.2 and 4.5, the stripping section is the region with the lower singularity values and therefore characterized by a good data assimilation capability. This suggests to injecting temperature-information in this part of the column.
However, the sensitivity analysis led to the question: is the passive structure the estimation scheme better suited for the ternary column considered? According to the constructive adjustable-estimation geometric approach (Alvarez, 2000, Lopez and Alvarez 2004), we know that the design of the estimation structure must be amount to a suitable compromise between reconstruction rate and robustness, depending on the estimation objectives, the model conditioning of the particular system, and the measurement uncertainty. Low estimation degrees favor robustness and disfavor the reconstruction rate. For this reason the passive structure, with maximum robustness, must be seen as the point of departure to draw estimation scheme with the best compromise between robustness and performance. Moreover, the detectability structure, with maximum estimation orders equal to the observability indices, constitutes the limit of performance in the absence of modeling errors. For this reason, the particular ill-conditioned temperature-composition relationship of the system suggests that the estimation structure better suited for the case study considered should be more on the passive side. The verification of this conjecture will be the object of the next Section.

4.4.2 The Passivated Structure

In a way that is analogous to the recursive robust control design via passivation (Krstic et al., 1995; Alvarez et al., 2005), and motivated by the decentralized control design for distillation columns (Castellanos-Sahagun et al., 2005), let us consider an estimation structure which is obtained by means the parallel combination of the two passive structures, equation (4.16a,b), presented in the previous Section, denoted as passivated structure.

The estimation structure with only one temperature sensor located in the generic i-th stage assumes the following form:

$$\Sigma_j = \{ \pi_j, x_j = ((c,^E,i), (c,^T,i)) \}$$

The observability matrix is characterized by the off diagonal terms equal to zero in order to prevent interaction between the matrix entries.

$$O(c,^i) = \begin{bmatrix} \beta cE(c,^i) & 0 & 0 \\ \beta cT(c,^i) & 0 & 0 \end{bmatrix}$$

The interactions between the two single-model estimation schemes are neglected in the same way by which a decentralized configuration control is design. It is important to note that this estimation scheme is characterized by a decentralized error propagation structure, with two passive error propagation mechanisms, one for each innovated state, which has been already displayed in figures 4.2 and 4.5.
Let us conclude that through the sensitivity analysis and the thermodynamic consideration presented in Chapter 2 is obtained a feasible estimation structure combining two single-passive estimation structures. The sensor location assessment of the passive cases is inherited by the passivated structure presented in this Section.

The methodology developed permitted to individuate three-candidate estimation scheme that in the next Chapter will be tested.

4.4.3 Generalization

As remarked in the previous Section, the passivated estimation structure is characterized by the injection of the information content in a reduced set of innovated states considering: (i) one fixed temperature sensor and (ii) the open loop of the noninnovated states. The estimation scheme obtained can be implemented for each sensor with good data assimilation capability or equivalently, with the lower singularity measures. This implies an algorithm with a modular construction obtained combining more single-sensor single-innovated scheme and injecting information in few column stages per temperature measurements. In particular, the application of a modular structure may be useful when the column operates with large operating condition changes, as batch column. Furthermore, the batch operation mode presents one additional feature, which makes the state estimation a more challenging task: batch distillation is an intrinsically dynamic process. During the course of the entire operation, composition profiles and operating conditions can change over a wide range of values, and the soft sensor must be designed to deal with the time-varying nature of the batch column. Implementing by means of the combination of more single-sensor single-innovated scheme in the column region with a good data assimilation capability it is possible monitoring the product composition accounts the challenging time-varying nature of the column. This modular estimation scheme is able to interrupt the error propagation mechanism in those column regions where the error propagation dominates the information assimilation.

Let us conclude pointing out that an estimation structure can be designed by combining two or more low order single-sensor schemes for the innovated states and an open-loop estimator for the noninnovated state. This estimation structure is independent of the particular estimation algorithm chosen, and it is possible to adjust the structure (modular scheme) to efficient perform the data assimilation task following the time-varying column behavior.
4.5 Fixed versus adjustable structure

In this Section let us consider the possibility of an adaptive estimation scheme using the information content transmitted by all the available temperature measurements. Sometimes the choice to consider the minimum number of temperature sensor is dictated by the increasing complexity of the algorithm implementation (when the temperature sensors number increase) and the trade off between total observability and robustness. For this reason, in the previous Section the choice of a robust estimation scheme was obtained by means of a methodology based on the study of the temperature-composition relationship. Indeed, recalling the Figure 2.6 when the column moves close to the azeotrope condition the temperature derivatives with respect to the alcohol compositions, in the enriching section, are close to zero and the singularity measures, Figures 4.2 and 4.5, are higher than in the stripping section. This induces the ill conditioning of the observability matrix during the column transient, and it affects the estimate quality. By means of the sensitivity analysis it was possible to analyze the data assimilation mechanism along the column and to individuate the column region with the best data assimilation capability where inject information. However, the singularity measurements evolution show that the information content transmitted by means of the temperature sensor is not always significative and the relationship temperature-composition becomes weaker in the enriching section during the process separation evolution. This suggests that to locate a fixed temperature sensor in this region could not be a good choice considering that it does not provide useful information for product composition estimation. For this reason, let us consider the possibility to implement an adaptive estimation structure where the temperature sensor is moved according with the singularity measure evolution. The resulting adaptive adjustable-structure could be obtained connecting and disconnecting two single-sensor estimation schemes, equation (4.21), one in the stripping section, and one in the enriching section, on the basis of the evolution of the temperature wave along the column. This implies to change the selected temperature sensor, and consequently the subset of the innovated states following the maximum temperature gradient along the column. This methodology could be particularly useful either for column subject to large operating conditions and for batch column characterized by an intrinsically dynamic behavior.

Finally, defined the estimation structure, the next objective will be to test the algorithm that must be applied.
Part II
Structure-Algorithm Testing
Chapter 5

The Soft Sensor Design

*In this Chapter, the resulting data assimilation schemes proposed are implemented with the NGE and CEKF algorithms. Two low-order estimators are presented. In particular, a PI-NGE and an extended Kalman filter with a reduced injection mechanism (GEKF) are designed. Furthermore, an estimator characterized by an adaptive-adjustable structure, where the temperature sensor location is changed following the temperature wave, is developed.*
5.1 Introduction

In Chapter 3 we mentioned that in literature there are several nonlinear estimation techniques, and in this Thesis two algorithms have been considered: the Nonlinear Geometric Estimator (NGE) and the Conventional Extended Kalman Filter (CEKF); detailed information on the Kalman Filter can be found elsewhere (Gelb, 1988, Haykin, 2001), here and in Chapter 3 have been presented only the key steps.

Undoubtedly, the CEKF is the most widely used estimation technique. It has been successfully tested in continuous and batch systems over a wide range of separations and operating conditions. This technique is formally intended for completely observable system and it is characterized by an implementation that requires the tuning of the error covariance matrices, $Q$ and $R$, via trial-and-error (Oisiovici and Cruz, 2000; Venkateswarlu and Avantika, 2001) or optimization-based searches (Baratti et al. 1995, 1998), and the on-line integration of a set of auxiliary ordinary differential equations (ODEs) whose number grows rapidly with stage number and components. Indeed, when the CEKF is designed for a system with $n$ balances, the number of Riccati equations (that grows approximately with $n$ square) and the number of the parameters, that must be tuned, make a large computational burden on the Kalman solution. Moreover, it is not clear how the EKF nonlinearity and complexity features can be reconciled with the linearity and simplicity of the majority of the industrial linear (MIMO, decentralized, or one way-decoupled) PI and (MIMO) model predictive control schemes.

The point of departure in this Thesis was to develop a robustly convergent soft sensor based on the algorithm proposed by Alvarez and Lopez (1999). The geometric algorithm has been successfully applied to solve the local nonlinear estimation problem of a free radical homopolymer (Alvarez and Lopez, 1999; Alvarez and Lopez, 2000, Lopez, 2000), to infer the concentration in a catalytic reactor (Lopez et al., 2002), to monitor the pollutant concentration in an activated sludge process for wastewater treatment (Lopez et al., 2004), and to infer product composition in a binary distillation column (Tronci et al., 2005, Fernandez 2006). The algorithm is applicable to observable and detectable systems and is characterized by a systematic construction-tuning procedure. The negative aspect related to the development of a nonlinear geometric estimator for a distillation column is due to the coordinate transformation on which the estimator algorithm is based. The geometric approach requires a quasi linearization of the nonlinear system, which is obtained applying a Lie derivative transformation to the measured outputs, with a derivative order depending on the dimension of the state vector and
on the number of available measured outputs. It emerges that the design of a full order soft sensor for a system characterized by a high dimension of the estimate state vector requires the calculation of the Lie derivative up to the order of the state variable minus one. This involves a high order temperature derivative weakening the soft sensor robustness. At this point, it is necessary to underline the importance of the estimation structure choice selection introduced in Chapter 4. Indeed, the choice of the structure permits to understand if all the observable state variables or only a small set of them (partial innovation) must be updated without compromising the soft sensor performance. These properties make the nonlinear geometric estimator (NGE) of Alvarez and Lopez a valuable alternative when compared with other estimation approaches, such as the CEKF. Indeed, the geometric approach permits to circumvent the on-line integration of the Riccati equations. It also enables the consideration of the sensor locations and the innovated states as design degrees of freedom. This approach can be seen in previous polymerization reactors (Lopez and Alvarez, 2004) and (fixed-sensor) binary distillation column (Tronci et al., 2005) estimation studies.

In this Chapter, let us demonstrate that the choice of the estimation structure is independent from the algorithm chosen to infer the column composition. Moreover, let us show that the same performance obtained with a conventional CEKF (Baratti et al., 1998), characterized by a complete innovation mechanism, can also be provided with a partial innovation mechanism and with simple tuning criteria.

5.2 The Soft Sensor Implementation

5.2.1 The Nonlinear Geometric Algorithm (NGE)

In this Section the results related to the PI-NGE functioning associated to the estimation structures ($\Sigma_1, \Sigma_2, \Sigma_3$), equations (4.15a)-(4.16)-(4.21), are showed. A PI-NGE was adopted in order to obtain more tolerance to modeling and measurements errors.

5.2.2 Construction and Tuning

Selected the estimation structure ($\Sigma_1, \Sigma_2, \Sigma_3$), the PI-NGE algorithm construction is obtained by the application of Theorem 3.4 (Section 3.3.4). According with the tuning scheme presented in Chapter 3, the soft sensor reference gains depend on the characteristic frequency ($\omega$) and on the damping factor ($\zeta$). Due to the almost linear output error dynamics, equation (3.70), the tuning of the pair ($\omega$, $\zeta$) can be performed according to conventional-like techniques and tuning guidelines for linear
filters (Alvarez and Lopez, 1999). The damping factor, $\zeta$, has been set equal to 1.5, in order to avoid an oscillatory response, for all the estimation structures considered, while the reference frequency has been chosen on the basis of the column dynamics. Considering the column dynamic response in the stripping section equal to a settling time of $t_a = 400$ min, and the characteristic time ($t_c = t_a/4$) $t_c = 100$ min, it was possible to calculate the corresponding natural characteristic frequency of the column in the stripping section ($\omega_h$).

$$\omega_h = 1/(100) = 10^{-2}$$  \hfill (5.1a)

Typically, $\omega_h$ is from 3-to-15 times larger (faster) than the natural frequency of the measurement response:

$$\omega_h = \omega_0 s_0$$ \hfill (5.1b)

then considering $s_0$ equal to 3, the characteristic frequency in the stripping section has been set equal to

$$\omega_h = 3 \times 10^{-2}$$ \hfill (5.1c)

In the enriching section, on the basis of the thermodynamic considerations introduced in Chapter 2, the characteristic frequency, $\omega_e$, was set equal to the ratio of the temperature derivatives in the enriching ($\beta',e$) and in the stripping section ($\beta',s$) by $\omega_e$:

$$\omega_e = \omega_h (\beta',e/\beta',s)$$ \hfill (5.1d)

$$\omega_e \sim 10^{-3}. \quad (5.1e)$$

The resulting gains have been adopted to tune the PI-NGE soft sensors without any further parameters calibration.

### 5.2.3 The Coupled Soft Sensor

Let us consider the structure $\Sigma_i$, equation (4.15a), where both ethanol and tert-butanol concentrations are innovated states with one sensor in the same stage. The corresponding soft sensor construction, for a sensor located in the generic $i$-th stage, has the following form (Alvarez and Lopez, 1999):

$$c^{'},i = f_i(c^{'},e_i, c^{'},t_i, c^{'},e_{i+1}, c^{'},t_{i+1}, u) + O_i^{'}(c^{'},e_i, c^{'},t_i, c^{'},e_{i+1}, c^{'},t_{i+1}, u)/\pi w + k_p[y_j - \beta(c^{'},j)] \quad (5.2a)$$

$$w_i = k_w[y_j - \beta(c^{'},j)] \quad c^{'},i = (c^{'},e_i, c^{'},t_i) \quad (5.2b)$$

$$c^{'},j = f_j(c^{'},e_j, c^{'},t_j, c^{'},e_{j+1}, c^{'},t_{j+1}, c^{'},e_{j-1}, c^{'},t_{j-1}, u), j \neq i, j \in [1,N] \quad (5.2c)$$
where the observability matrix $O$, equation (5.2d), as already stated in Chapter 4, can be considered as made up of two row vector. The entries of the first row vector are the first temperature derivatives with respect to ethanol and tert-butanol compositions, while the entries of the second row vector are characterized by the product of first or second order temperature derivatives by balance equations [equations (4.15c,d)]. The contribution due to the second order temperature derivative has been neglected in order to avoid the numerical error propagation. By means this approximation, the observability matrix, the estimability map and the tuning parameters ($k_p, k_w$) have the following expressions:

$$
O(c_i, c_{i-1}, c_{i+1}, u) = \partial_c \phi(c_i, c_{i-1}, c_{i+1}, u), \pi = (0, 1)'
$$

(5.2d)

$$
\phi(c_i, c_{i-1}, c_{i+1}, u) = \{\beta(c_i), [\partial_c \beta(c_i)]f(c_i, c_{i-1}, c_{i+1}, u)\}'
$$

(5.2e)

$$
k_p = (2\zeta + 1)(\omega, \omega')', k_w = \omega^j, u = (F, R, V, c_{i\beta})'
$$

(5.2f)

According to the sensitivity analysis presented in Chapter 4, Figure 4.1, the temperature sensor has been located in the column bottom (stage 0), which is the column region that offers the best means of effective data assimilation. The results of the soft sensor implementation related to the estimation structure $\Sigma_1$ are depicted in Figure 5.1 for the distillate and the bottom compositions. As it is possible to observe the combination between the estimation structure $\Sigma_1$ and the algorithm led to design a soft sensor that is not able to reconstruct the distillate and bottom composition behavior.

![Figure 5.1. The PI-NGE with a coupled estimation structure $\Sigma_1$.](image)

In Figure 5.2 is showed the condition number behavior in updated conditions. As expected, see Figure 4.1, the estimator malfunctioning is justified by the increasing error propagation along the column during the process separation.
As mentioned in Chapter 4, the reason of the soft sensor malfunctioning must be seen in the high interactions between the observability matrix entries, equation (5.2d). Even if the presence of the second order temperature derivatives has been neglected, the high ill conditioning of the observability matrix in the column top compromise the ethanol and tert-butanol estimate.

In conclusion, let us confirm the considerations already stated in Chapter 4. Definitely, the estimation structure $\Sigma_1$ is not suited for this column and the next step will be to test the other estimation structures with the NGE algorithm.

### 5.2.4 The single-sensor single-innovated state

In this Section the soft sensor functioning is tested with the estimation structures $(\Sigma_2)$, equations (4.16b)-(4.16c), previously introduced in Chapter 4. Let us recall the two estimation structures:

\[
\Sigma_2 = (\Sigma_{2a}, \Sigma_{2b})
\]

\[
\Sigma_{2a} = \{ \pi_{ij}, x_I = (c_i, {^E}c_i), x_{II} = (c_i, {^T}c_i) \}
\]

\[
\Sigma_{2b} = \{ \pi_{ij}, x_I = (c_i, {^T}c_i), x_{II} = (c_i, {^E}c_i) \}
\]

where with $\Sigma_{2a}$ is denoted the structure with ethanol as innovated state, and with $\Sigma_{2b}$ the structure where tert-butanol is the innovated state.
**Ethanol as innovated state**

The combination of the estimation structures $\Sigma_{2a}$, equation 5.3b, with the PI-NGE algorithm led to the following equations:

\[
\hat{c}_{\text{E},i} = f_E(c_{\text{E},i}, c_{\text{T},i}, c_{\text{E},i-1}, c_{\text{T},i-1}, c_{\text{E},i+1}, c_{\text{T},i+1}) + \left\{ \frac{1}{\beta_E} \right\} \left[ w + 2\zeta_0 (y_i - \beta(c_{\text{E},i}, c_{\text{T},i})) \right]
\]

(5.4a)

\[
w_i = \omega^2 (y_i - \beta(c_{\text{E},i}, c_{\text{T},i}))
\]

(5.4b)

\[
\hat{c}_{\text{T},i} = f_T(c_{\text{E},i}, c_{\text{T},i}, c_{\text{E},i-1}, c_{\text{T},i-1}, c_{\text{E},i+1}, c_{\text{T},i+1}), k = E, T
\]

(5.4c)

\[
\hat{c}_{\text{k},j} = f_k(c_{\text{E},j}, c_{\text{T},j}, c_{\text{E},j-1}, c_{\text{T},j-1}, c_{\text{E},j+1}, c_{\text{T},j+1}), j \neq i, j \in [1,N]
\]

(5.4d)

where $\omega_E, \omega_T, \zeta_E, \zeta_T$ are the tuning parameters; $w$ represents the integral action defined as a dynamical state that estimates and compensates the effect of modeling errors in the predicted output, or equivalently, eliminates the output error mismatch.

According to the sensitivity analysis, Section 4.4.1, and the thermodynamic consideration made in Section 2.4, the stripping section is the column region with the least error propagation. This is the reason for which the temperature sensor is located in the column bottom (stage 0). However, it is important to consider that, during the process evolution, especially in the middle of the column the difference between the sensitivity values reached in the enriching and in the stripping section is not marked. This suggests that a sensor in this column region could be selected in order to find a compromise between robustness and closeness to the effluent compositions. This means that if our objective is to estimate two product compositions (distillate and bottoms) two are the possible temperature sensor locations. Consequently these configuration schemes must be considered:

\[
\Sigma_{2a} = \{ \pi_{1,0}, x_i = (c_{\text{E},i}), x_{II} = (c_{\text{T},i}) \}
\]

(5.5a)

\[
\Sigma_{2b} = \{ \pi_{1,18}, x_i = (c_{\text{E},i}), x_{II} = (c_{\text{T},i}) \}
\]

(5.5b)

The two estimation structures are characterized by having the same number of sensors and the same innovate/noninnovated state set but different temperature sensor locations. In the first estimation structure (5.5a) the sensor is located in the column bottom (stage 0), while in the estimation scheme (5.5b) only one sensor is located in the stage 18. In Figure 5.3 is presented the comparison between the model predictions, the PI-NGE algorithm, and the off-line experimental when the sensor is located in the column bottom (stage 0) and it is referred to the Run I. As predicted by the sensitivity plot of Figure 4.2, the column bottom exhibits a good data
assimilation capability. Indeed, it is possible to observe that the agreement between estimated and off-line experimental data is quite good in the column bottom for the ethanol composition, while the same cannot be stated for the tert-butanol component.

In the enriching section the column model performs better than the estimator. This could be justified by the absence of information content injected in the enriching section.

In Figure 5.4 is showed the comparison between the model predictions, the PI-NGE algorithm and the off-line experimental data when a sensor is located in the enriching section (stage 18). When the sensor is located in the enriching section the soft sensor performances are similar to the model predictions. This means that only
one temperature sensor in the enriching section is not able to reconstruct the state trajectories. The reasons must be seen in the weak temperature-composition relationship that characterizes the enriching section during the process separation evolution. Finally, here are not represented but the same results hold when the soft sensor is tested with the experimental data referred to the Run II (Table 2.4).

**Tert-butanol as innovated state**

When the tert-butanol is the innovated state, equation (5.3c), and the sensor is located at the (variable) i-th stage, the combination of the estimation structure and the algorithms becomes:

\[
c_{i,j}^e = f_{i,j}^e(c_{i-1,j}^e, c_{i,j}^T, c_{i,j}^c, c_{i+1,j}^e, c_{i+1,j}^T) \\
\]

\[
c_{i,j}^T = f_{i,j}^T(c_{i-1,j}^e, c_{i,j}^T, c_{i,j}^c, c_{i+1,j}^e, c_{i+1,j}^T) + \frac{1}{\beta(c_{i,j}^e, c_{i,j}^T)} \{ w + 2\zeta\omega[y_j - \beta(c_{i,j}^e, c_{i,j}^T)] \} \\
\]

\[
w_j = \omega^2[y_j - \beta(c_{i,j}^e, c_{i,j}^T)] \\
\]

\[
c_{i,j}^k = f_{i,j}^k(c_{i,j}^e, c_{i,j}^T, c_{i,j}^c, c_{i,j}+p^e, c_{i,j}^T, c_{i,j}+p^T), j \neq i, j \in [1,N] \\
\]

Recalling the sensitivity measure evolution showed in Figure 4.5 and the stability margin in Figure 4.7, the tert-butanol should definitely not be chosen as the innovated state when the measurement is located in the stage interval [20-30]. However, if a temperature sensor in the enriching section must be selected, it must be located in the stages [10-18]. As in the previous case, equation 5.5, two estimation structures can be considered:

\[
\Sigma_{2a} = \{\pi_{1,0}, x_1 = (c_{1}^T), x_2 = (c_{1}^e)\} \quad (5.7a) \\
\Sigma_{2b} = \{\pi_{1,18}, x_1 = (c_{18}^T), x_2 = (c_{18}^e)\} \quad (5.7b) \\
\]

In the Figure 5.5 is showed the comparison between PI-NGE algorithm the model predictions and the off-line experimental data referred to the Run I. The estimator performance is similar to the ones of the case with ethanol as innovated state, Figure 5.3, and with only one temperature sensor in the column bottom. It is important to observe that if the objective is to estimate the ethanol and the tert-butanol compositions in the column bottom then it is better to locate the sensor in the column bottom, while if the objective is to estimate the column top then probably the temperature sensor should be located in the stage 18. In the Figure 5.6 it is showed the comparison between the PI-NGE outputs, the model predictions and the off-line experimental data when the sensor is located in the stage 18. It is possible to observe that locating the temperature sensor in the enriching section the estimate of
the top composition improves if compared with the Figure 5.5, but the estimator performance are similar to the model predictions. As in the previous case, Figure 5.4, this is related to the weak assimilation data capability of the enriching section.

Figure 5.5. Single-sensor (column bottom) passive estimation (tert-butanol composition as innovated state).

The same results hold when the soft sensor is tested with the experimental data referred to the Run II. This confirms that a passive structure is not able to reconstruct both the product composition in both the enriching and in the stripping section.

Figure 5.6. Single-sensor (stage 18) passive estimation (tert-butanol composition as innovated state).
Remarks

Let us conclude this Section with some consideration concerning the soft sensor design by means of the implementation of a passive estimation structure, equations 5.5-5.7. From the Figures 5.3, 5.4, 5.5, 5.6 it is possible to note that a single-sensor single-innovated estimation scheme is able to reconstruct the dynamic of the innovated state only in the column region where the temperature sensor is located. In particular when the sensor is located in the column bottom, which correspond to the region with a good data assimilation mechanism, the agreement between inferred and experimental data is rather good, while the same consideration cannot be stated if the sensor is located in the enriching section. This is the reason for which if both the product composition must be estimated it is necessary to design a soft sensor with a robust estimation structure. In the next Section let us test the estimator performance when the passivated structure, equation (4.21), is implemented.

5.2.5 The decoupled Soft Sensor

After having analyzed the passive soft sensors, \((\Sigma A, \Sigma b)\), our objective is understand if it is possible to improve the soft sensor performance. The combination of the estimation structures \(\Sigma i\), equation 4.21, jointly with the sensitivity analysis presented in Chapter 4, and with the PI-NGE algorithm leads to the design of a decoupling estimator characterized by the following equations:

\[
\begin{align*}
c_{E,i} & = f_E(c_{E,i}, c_{T,i}, c_{E,i+1}, c_{T,i+1}) + \frac{1}{\beta_{cE}(c_{E,i}, c_{T,i})} \left\{ w_E + \frac{2}{\omega_E} [y_i - \beta_{cE}(c_{E,i}, c_{T,i})] \right\} \\
c_{T,i} & = f_T(c_{E,i}, c_{T,i}, c_{E,i-1}, c_{T,i-1}, c_{E,i+1}, c_{T,i+1}) + \frac{1}{\beta_{cT}(c_{E,i}, c_{T,i})} \left\{ w_T + \frac{2}{\omega_T} [y_i - \beta_{cT}(c_{E,i}, c_{T,i})] \right\} \\
c_{k,i} & = f_k(c_{E,j}, c_{T,j}, c_{E,j+1}, c_{T,j+1}), \quad j \neq i, j \in [1, N] \\
w_{E} & = \omega_E^2 [y_i - \beta_{cE}(c_{E,i}, c_{T,i})], \quad w_{T} = \omega_T^2 [y_i - \beta_{cT}(c_{E,i}, c_{T,i})]
\end{align*}
\]

The analysis of the singularity measures, Figures 4.2 and 4.5, and of the noninnovated dynamics, Figures 4.4 and 4.7, suggests that the stripping section is a convenient region where to inject information. In the Figures 5.7 and 5.8 is shown the comparison between the PI-decoupled NGE performance and the off-line experimental data referred to for two different experiments (Run I and Run II).
As it is possible to observe in Figures 5.7 and 5.8 the agreement between the soft sensor performance and the experimental data is rather good both in the enriching and in the stripping sections.

The soft sensor designed is characterized by decentralized error propagation. In Figure 5.9 is showed the sensitivity measure referred to ethanol and tert-butanol component, when the sensor is located in the column bottom (stage 0), in updated condition. It is important to remark that the sensitivity measure show the same behavior for both the component and assume small values during the process evolution.
A fixed temperature sensor was added in the enriching section, and as can be observed in Figure 5.10 the estimate top composition poorly improves. In this case, the results are obtained implementing two passivated structures, one in the stripping section where the information is injected in the column bottom (stage 0), and the other in the enriching section where, according to the sensitivity analysis, the temperature sensor is located in the tray 18. However, injecting information content in the enriching section, which is the column region characterized by a weak relationship temperature-composition, the distillate estimate improvement is marginal so that a passivated module in the enriching section is not justified.

Let us conclude that in the column considered in this Thesis the presence of a passivated structure in the enriching section is not necessary for the purpose of...
monitoring the product compositions. In particular this is related to the thermodynamic considerations presented in Chapter 2. Finally, it is important to remark that a PI-NGE with only 66 differential equations (64 balance equations and two integral actions) and four tuning parameters (ωₜ, ωₗ, ζₑ, ζ₉) is able to reconstruct top and bottom dynamics.

5.3 Extended Kalman Filter with reduced data injection

In this Section, let us demonstrate that the estimation scheme $\Sigma_i$, equation (4.21), adopted to developing a non linear geometric algorithm can be applied to the EKF obtaining a reduced EKF that let us denote as (GEKF) to distinguish it from the CEKF.

\[
\begin{align*}
\hat{c}_i^{E} &= f^E_i(c_i^{E}, c_i^{T}, c_{i-1}^{E}, c_{i-1}^{T}, c_{i+1}^{E}, c_{i+1}^{T}) + \left[ \frac{\beta c_i^{E}(c_i^{E}, c_i^{T})}{\sigma_{11}} \right] \frac{1}{r} \sum_j (c_j^{E} - c_i^{E})^2 \tag{5.9a} \\
\hat{c}_i^{T} &= f^T_i(c_i^{E}, c_i^{T}, c_{i-1}^{E}, c_{i-1}^{T}, c_{i+1}^{E}, c_{i+1}^{T}) + \left[ \frac{\beta c_i^{T}(c_i^{E}, c_i^{T})}{\sigma_{22}} \right] \frac{1}{r} \sum_j (c_j^{T} - c_i^{T})^2 \tag{5.9b} \\
\hat{c}_i^{k} &= f^k_i(c_i^{E}, c_i^{T}, c_{j-1}^{E}, c_{j-1}^{T}, c_{j+1}^{E}, c_{j+1}^{T}), j \neq i, j \in [1, N] \tag{5.9c}
\end{align*}
\]

\[R = r = 1.67E-03\]

From the structural point of view, the GEKF is characterized by only 3 Riccati equations, 3 tuning parameters obtained by means of an optimizing technique (Baratti et al., 1998), and a diagonal model error covariance matrix $Q$, equation (3.27). In Figures 5.11 and 5.12 is showed the comparison between the GEKF and the off line experimental data referred to the experimental tests, Run I and Run II, represented in Table 2.4 and considering only one temperature sensor located in the column bottom.

![Figure 5.11. Comparison between the GEKF and off-line experimental data referred to Run I.](image)
The advantages of the GEKF over the CEKF reside in the fact that the construction implementation of the GEKF is simpler if compared with a CEKF. Indeed, the GEKF algorithms are characterized respectively by 64 nonlinear model ODE’s and an additional reduced subset of structural equations.

5.4 Comparison between GE, GEKF and CEKF

Here, the PI-NGE and GEKF estimators with one temperature sensor (column bottom) are compared with a conventional CEKF with two sensors (in the column bottom and stage 26), as it is generally done in distillation column studies. The resulting behavior is presented in Figure 5.13 and Figure 5.14, showing that basically the PI-NGE, equations (5.8), and the GEKF, equations (5.9), yield the same performance of the CEKF.
It is important to note that the algorithms developed with the decoupled model estimation give the same performance of the EKF complete characterized by 2144 ODE’S and six tuning parameters. In other words, the complete EKF with 2144 equations really has redundancy in the sense that a full profile data mechanism is set and on-line solved, but the information is effectively utilized only in a small column section, about stage \([0 – 7]\).

In conclusion, when a robust estimation structure is obtained then the data assimilation mechanism of the innovated subsystem can be designed either with NGE or GEKF. The results obtained justify the main objective of this Thesis: to develop a robust estimation structure independent from the algorithm chosen to tackle the estimation task.

### 5.5 The Adaptive-Adjustable Soft Sensor

Considering that good bottom composition estimates are obtained with one fixed temperature sensor located in the column bottom (stage 0), in this Section let us analyze the possibility of extracting information content from all the available temperature measurements in order to assess the convenience of adapting the measurement location during the time along the column. As it can be seen in Figures 4.1, 4.2, 4.6, the temperature front changes with time, and this suggest moving the sensors according to the temperature wave front. The resulting adaptive adjustable-structure is obtained connecting and disconnecting two passivated structure, equations (4.21), one in the stripping section and one in the enriching section, on the basis of the evolution of the temperature wave along the column. The adjustable
estimation structure has been combined with the P-NGE characterized by the following equations:

\[
c_{E,i} = f^E_i(c_{E,i}, c_{E,i+1}, c_{E,i-1}, c_{T,i+1}, c_{T,i-1}, c_{E,i}, c_{T,i}) + \{1/\beta c_E(c_{E,i}, c_{T,i})\} \{\omega E[yi - \beta(c_{E,i}, c_{T,i})] \}
\]

(5.8a)

\[
c_{T,i} = f^T_i(c_{E,i}, c_{T,i+1}, c_{E,i+1}, c_{E,i-1}, c_{T,i-1}, c_{E,i}, c_{T,i}) + \{1/\beta c_T(c_{E,i}, c_{T,i})\} \{\omega T[yi - \beta(c_{E,i}, c_{T,i})] \}
\]

(5.8b)

\[
c_{k,j} = f^k_j(c_{E,j}, c_{T,j}, c_{E,j+1}, c_{E,j-1}, c_{T,j+1}, c_{T,j-1}, c_{E,j}, c_{T,j}), j \neq i, j \in [1, N]
\]

(5.8c)

The reason for which a P-NGE has been adopted in place of a PI-NGE is to prevent the phenomena of reset-windup. The selected temperature sensor was changed, online, and was consequently changed the subset of the innovated states (VLE relationship). Specifically, two temperature sensors are considered: one in the stripping and one in the enriching section. The selected sensors are changed in the enriching section among the stages \{12, 18, 22, 26, 30\} and in the stages \{0, 4\} in the stripping section. The temperature sensor was adapted, during the separation process, considering the evolution of maximum temperature gradient in the stripping and in the enriching section. Recalling the transient temperature profiles evolution along the column, see Figure 2.4, it is clear that after 50-60 minutes the temperature gradient in the enriching section is negligible (small) and assumes the high value in stage 12 close to the feed stage, while at the same time, in the stripping section, it moves from the stage 5 toward stage 0 (column bottom) and there it remains since the end of the process separation.

It is important to remark that when the maximum temperature front is located in a column stage where a temperature sensor is not available then the closer stage is selected.
The performance of the soft sensor with an adaptive adjustable-structure is compared with the performance of a PI-NGE with one fixed temperature sensor (located in the stage 0) and the behavior is showed in Figure 5.15 and Figure 5.15. In particular, the information content, injected in the enriching section, does not improve in a significative way the composition estimation in the distillate. This is due to the thermodynamic consideration previously done in Chapter 2. However the adaptive structure proposed could be particularly useful for batch columns intrinsically characterized by dynamic behavior.

5.6 Some Concluding Remarks

Let us conclude the soft sensor design Section with some considerations.
In Chapter 4 it has been presented a methodology to select the innovation mechanism based on a geometric approach. By means of this methodology it was possible to select a robust estimation structure applicable to any algorithms. This was demonstrated in the Chapter 5 where the estimation structure selected has been tested with the NGE and the GEKF algorithms. Moreover, the design of an estimator with an adaptive estimation structure has been tested. In particular, it is important to remark that the GEKF as the NGE are characterized by a reduced innovation mechanism and consequently by a systematic construction.
Chapter 6

The Adaptive Kalman Filter

In this Chapter, the problem of tuning the Conventional Extended Kalman Filter, CEKF, is solved by using the physical meaning of the model covariance matrix to obtain a proper and efficient estimator. The effectiveness of the method is tested on a binary distillation and the inferred compositions are compared with experimental data obtained in the pilot plant described in Chapter 2.
6.1 Introduction

The Conventional Extended Kalman Filter (CEKF) is by far the most widely used estimation technique. By mean this algorithm it is possible to take into account the nonlinearities of the system without loosing the recursive update form of the linear Kalman filter. The conditional probability distribution, for nonlinear system, is not Gaussian even if the noise processes and the initial error are assumed to be Gaussian. This means that to determine the distribution an infinite number of moments are required and the moments are coupled in increasing order. To obtain practical estimation algorithm, methods to compute the mean and the covariance that did not depend upon knowing the probability density function are needed (Gelb, 1988; Haykin, 2001; Welch and Bishop, 2001; Muske and Edgar, 1977).

Several studies, designed to improve this estimation algorithm have been reported in literature. Moghaddamjo and Kirlin (1989) proposed a Kalman Filter modified to follow the changes in input forcing functions and noise statistics. Dimitratos (1989, 1991) applied an adaptive CEKF, which employs continuous-time process models for dynamics to estimate the time evolution of the reacting mixture composition in a polymerization reactor. Zhou and Luecke (1995) presented a procedure using observations with linear process to develop estimates for the effective values of the covariance matrices of the process model and of the measurements errors. Julier and Uhlmann (1995) proposed the Unscented Kalman Filter (UKF) further developed by Wan and van der Merwe (2000). The UKF is characterize to have a state distribution approximated with the GRV (Gaussian Random Variable), this is represented using a minimal set of chosen sample point which are able to capture the true mean and covariance of the GRV (Haykin, 2001). Venkateswarlu and Avantika (2001) applied a continuous version of the Adaptive Fading Kalman Filter (AFKF) with the objective to infer composition in batch distillation columns. The computational procedure of this algorithm is similar to that of the CEKF except the covariance of the prediction equation, which is characterized by a forgetting factor in order to form an adaptive covariance matrix.

The objective of the algorithms above mentioned is to improve nonlinear estimation. In order to be applied they need a significant set of experimental data and a time-consuming numerical optimizing procedure to tune the covariance matrices. In particular, the tuning of the entries of the filter covariance matrices relies on trial and error procedure to achieve maximum likelihood state estimation.

In this Chapter, let us propose a methodology to tune the CEKF based on an approximation of the model error covariance matrix (c.f., Leu and Baratti, 2000).
The measurements covariance matrix is assessed either from the information of the measurements instruments or from a standard test on the measured signals, while the model error covariance is evaluated considering the uncertainty of the model parameters set. The methodology is applied to monitoring the outlet composition in binary distillation columns, and it is tested using the same experimental data presented in Chapter 2.

6.2 The Covariance Matrices of the Estimates

*How much can we change the estimates and fit the data well?* (Bard, 1974)

As already mentioned in the previous Section the procedure to tune the CEKF is based on the approximation of the covariance matrix. In particular, it was assumed the model errors only due to the imprecision of model parameters. In this Section a procedure to evaluate the experimental covariance matrix, \( Q_p \), is presented more details can be found in Bard (1974).

Let us indicate the dependence of the objective function \( \Phi \) from parameters and measured values \( w \) by writing \( \Phi(p, w) \). At the minimum we have (Bard, 1974):

\[
\partial \phi(p^*, w)/\partial \theta = 0 \tag{6.1}
\]

Varying the data slightly then the minimum point will shift:

\[
\partial \phi(p^* + \delta p^*, w + \delta w)/\partial \theta = 0 \tag{6.2}
\]

Expanding equation (6.2) in Taylor series and considering only terms up to first order we can obtain

\[
(\delta^2 \phi/\delta p^2)\delta p^* + (\delta^2 \phi/\delta p \delta w)\delta w \approx 0 \tag{6.3}
\]

\[
\delta p^* = -G^{-1}(\delta^2 \phi/\delta p \delta w)\delta w \tag{6.4}
\]

where \( G^* = \delta^2 \phi/\delta p^2 \bigg|_{p=0^*} \) is the Hessian of the objective function with respect to the parameters (calculated at the minimum).

The covariance matrix \( Q_p \) is defined by

\[
Q_p = E(\delta p^* \delta p^{T*}) \tag{6.5}
\]

so that

\[
Q_p \approx E(G^{-1}(\delta^2 \phi/\delta p \delta w)\delta w \delta w^T(\delta^2 \phi/\delta p \delta w)^T G^{-1}) \tag{6.6}
\]
The quantities $G^*$ and $\partial^2 \phi / \partial p \partial w$ are evaluated at $p = p^*$ and at the actual sample $w$. These quantities are constants, and can be written as

$$Q_p \approx G^{*-1} (\partial^2 \phi / \partial p \partial w) Q_w (\partial^2 \phi / \partial p \partial w)^T G^{*-1}$$

(6.7)

where $Q_w$ is the covariance matrix of the data

$$Q_w \equiv E(\delta w, \delta w^T)$$

(6.8)

Assuming that $w_\mu$ has a covariance matrix $Q_\mu$ and is independent of $w_\eta$ ($\eta \neq \mu$) then equation (6.7) reduces to:

$$Q_p \approx G^{*-1} \left[ \sum_{\mu=1}^N \left( \frac{\partial^2 \phi}{\partial p \partial w_\mu} \right) Q_\mu \left( \frac{\partial^2 \phi}{\partial p \partial w_\mu} \right)^T \right] G^{*-1}$$

(6.9)

Specific results can be obtained when the objective function depends on the moment residual matrix $M$. This class of functions (sums of squares, log-likelihood for normal distributions) admits the following Gauss approximation for $G$.

$$G \approx 2 \sum_{\mu=1}^N \Gamma_\mu B_\mu$$

(6.10)

where

$$B_\mu \equiv -\partial c_\mu / \partial p = \partial f_\mu / \partial p, \quad \Psi'(M(p)) \equiv \phi(p), \quad \Gamma \equiv \partial \Psi / \partial M$$

(6.11)

For reduced model with $w_\mu = y_\mu$

$$\partial^2 \phi / \partial p \partial y_\eta \approx -2 B_\eta^T \Gamma$$

(6.12)

Substituting equation (6.10) and equation (6.12) in equation (6.9) is obtained:

$$Q_p = \left( \sum_{\mu=1}^N B_\mu \Gamma B_\mu \right)^{-1} \left( \sum_{\mu=1}^N B_\mu \Gamma V_\mu \Gamma B_\mu \right) \left( \sum_{\mu=1}^N B_\mu \Gamma B_\mu \right)^{-1}$$

(6.13)

For a single equation least-squares, where $\sigma$ is the standard deviation of the observations, we have

$$B_\mu = b_\mu = \partial f_\mu / \partial p, \quad \Gamma = 1, \quad Q_\mu = \sigma^2$$

(6.14)

the equation (6.13) become:
\textbf{6.3 The a priori CEKF}

In this Section the CEKF characterized by a systematic tuning procedure is presented. In order to present the characteristic equations of the a priori EKF, let us recall the equations of a general nonlinear continuous dynamics system:

\begin{align}
\dot{x}(t) &= f(x,u) + w(t), \quad x(0) = x_0 + w_0 \\
y(t) &= h(x,u) + v(t)
\end{align}

(6.16) (6.17)

where equation (6.16) represents the state equation, and equation (6.17) the relationships between the outputs \(y\) and the states \(x\). The process noises \(w(t)\), \(v(t)\) as well as the initial error \(w_0\) are assumed to have zero mean and Gaussian distributions.

For convenience let us recall the equations governing the continuous-discrete CEKF (Ray, 1981; Gelb 1988, Haykin, 2001):

- **State estimation and error covariance propagation (Riccati equations)**

\begin{align}
\dot{x} &= f(\hat{x}(t), u) \\
\dot{P}(t) &= F(\hat{x}(t), u)P(t) + P(t)F^T(\hat{x}(t), u) + Q(t) \\
x(t_0) &= x_0 \quad \text{and} \quad P(t_0) = P_0
\end{align}

(6.18) (6.19)
• **Update**

\[
\hat{x}_k = \hat{x}_k^- + G_k \left[ y_k(t) - h_k(\hat{x}_k^-) \right],
\]

\[
P_k = \left[ I - G_k H_k(\hat{x}_k^-) \right] P_k^-
\]

with the gain matrix defined as:

\[
G_k = P_k^- H_k^T(\hat{x}_k^-) H_k(\hat{x}_k^-) P_k^- + R \right)^{-1}
\]

where \( \hat{x}(t) \) is the estimated state vector, \( Q(t) \) is the covariance matrix of the process noise and \( R \) is the covariance matrix of the measurements errors. The matrices \( F_k \) and \( H_k \) are the Jacobian with respect to the state vector, evaluated in the estimated values:

\[
F_k(\hat{x}(t),u) = \frac{\partial f(\hat{x}(t),u)}{\partial \hat{x}(t)} \bigg|_{\hat{x}(t)=\hat{x}(t)}
\]

\[
H_k(\hat{x}(t)) = \frac{\partial h_k(\hat{x}(t))}{\partial \hat{x}(t)} \bigg|_{\hat{x}(t)=\hat{x}(t)}
\]

In the equation (6.21) it is possible to observe that the matrix \( G(t) \) corrects the model prediction using the innovation-error difference between the actual and predicted outputs, by weighting the model prediction and the actual measurement contributions by mean of \( Q \) (model covariance matrix) and \( R \) (measure covariance matrix). Generally, the error covariance matrices are parameters that must be tuned to obtain the best performance of the filter. Usually, this purpose is achieved carrying out dynamic experiments and a trial-and-error procedure is usually implemented (Oisiovici and Cruz, 2000; Welch and Bishop, 2001; Baratti et al., 1998). This method may be time-consuming and it is one of the main limit of this estimator. However, physical considerations can be used in order to assess the experimental model error covariance matrix by means the approximation illustrated in the previous Section (equation 6.15).

Assuming that: (i) the uncertainty arises only from the model parameter uncertainties; (ii) the model of the plant consists of the nominal model with a first-order Taylor series expansion with respect to the parameters; (iii) the parameter errors have white noise with zero mean, and have normal distribution with a constant covariance matrix that can be determined, and possibly updated, from online data. Giving assumptions (i) and (ii) a mathematical form, the following
expression for the model described by the equation (6.16) can be derived (Leu and Baratti, 2000):

$$\dot{x} = f(x,u,p) + F_p(x,u)(p - \hat{p}) + o(x,u,p - \hat{p})$$  \hspace{1cm} (6.23)

where $p$ is the real unknown value of the parameter and $\hat{p}$ its estimated value. $F_p$ is the Jacobian with respect to the model parameter $p$. Thus, the new additive error is given by

$$w(x,u,p - \hat{p}) = F_p(x,u)(p - \hat{p})$$  \hspace{1cm} (6.24)

and the covariance matrix of the process noise becomes state and input-dependent

$$Q(x,u,Q_p) = F_p(x,u)Q_pF_p^T(x,u)$$  \hspace{1cm} (6.25)

once the covariance matrix of the model parameters is computed, it is possible to derive the covariance matrix of the model by using the equation (6.25), which substituted in the equation (6.18), gives the resulting Riccati equation:

$$\dot{P}(t) = F(\hat{x},u)P(t) + P(t)F_p^T(\hat{x},u) + F_p(\hat{x},u)Q_pF_p^T(\hat{x},u)$$  \hspace{1cm} (6.26)

The CEKF characterized by this approximation of the covariance matrices is denoted as a priori CEKF.

### 6.3.1 Results and discussion

The performance of the a priori CEKF has been assessed by comparison with the dynamic of the binary system previously described in Chapter 2.

To develop the a priori CEKF we need first to select the measured outputs among the available ones and then to construct the matrix $Q$, which implies to identify the model parameters, $p$, responsible of the modeling errors. According to previous study (Baratti et al., 1995) two-temperature sensors are adopted, one is located in the column bottom (stage 0) and one is located in the enriching section (stage 26). At this point to design the algorithm it is necessary to evaluate the matrix covariance error $Q$. Previous studies on this system (Tronci et al., 2005) evidenced that the vapor-liquid equilibrium strongly affects the estimation accuracy, and the selection of a good data set for the thermodynamic model is an important aspect in order to obtain good performance of the inferential system. This is because the equilibrium relationship, equation (6.17), relates the measured outputs with the states, and corrects them according to the thermodynamic model. Therefore, the proposed covariance matrix was obtained by calculating the Jacobian of the 32-order system of ordinary differential equations that reconstructs the dynamics of the liquid
composition, with respect to the three NRTL parameters: \( g_{12}, g_{21}, \alpha \) (Gmehling and Onken, 1977a,b). As a further simplification, \( Q_p \) was assumed a diagonal matrix, that is \( Q_p = \sigma I \), where \( I \) is the identity matrix and \( \sigma \) is the variance matrix.

The effectiveness of the a priori tuning was verified by considering two different sets of NRTL parameter. Set I (Gmehling and Onken, 1977a) is the more accurate for the operation range of the column, while Set II (Gmehling and Onken, 1977b) is characterized by a higher standard deviation.

<table>
<thead>
<tr>
<th></th>
<th>Set I</th>
<th>Set II</th>
</tr>
</thead>
<tbody>
<tr>
<td>( q_{p(i,j)} ) for ( i=j )</td>
<td>0.03</td>
<td>0.0072</td>
</tr>
<tr>
<td>( q_{p(i,j)} ) for ( i \neq j )</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>( r(1,1) )</td>
<td>0.29</td>
<td>0.29</td>
</tr>
<tr>
<td>( r(2,2) )</td>
<td>0.045</td>
<td>0.045</td>
</tr>
<tr>
<td>( r(1,2) = r(2,1) )</td>
<td>0.0</td>
<td>0.0</td>
</tr>
</tbody>
</table>

It is important to note that the entries of the measurements error covariance matrix, \( R \), coefficients are kept constant for the two set of parameters, because \( R \) only depends on the accuracy of the measurement instruments. The performance of the soft sensor is tested with experimental data obtained with two tests reported in the Table 2.5.

In the Figure 6.1 are represented the distillate and bottoms compositions referred to the Run I. In particular, the ethanol effluent compositions inferred with the a priori CEKF for two different sets of NRTL parameter is compared with the experimental data and the results obtained by simple model integration. It is worth noting that the data referred to the model are obtained using the more accurate set of NRTL parameters (Set I).

![Figure 6.1. Comparison between experimental data, a priori CEKF with Set I,](image-url)
The improvement due to the a priori CEKF is evident for both the sets of parameters in the bottom reconstruction, where the model is not able to follow the actual dynamics, because of its intrinsic inaccuracy (Baratti et al., 1995). The Figure 6.2 represents the results referred to the experimental Run II. Again, the estimator is able to correct the product composition obtained by integrating the simplified model of the column, being the improvement more evident in the bottom.

It is worth noting that the composition profiles inferred with the CEKF using the two different sets of parameter do not match perfectly. As already mentioned in Chapter 2, the column behavior is affected not only by the equilibrium parameters but also by the dynamic hydraulic parameters. This means that the disagreement between CEKF and off-line experimental data is due either to the simplified form of the matrix $Q_p$ but also to neglecting liquid dynamic hold-up. A more accurate evaluation of $Q_p$ considering the uncertainty due to the equilibrium and hydraulic parameters can improve the results. In any case, satisfying estimator performances are obtained just considering the standard deviation of the observations from which the NRTL parameters were calculated, while expensive and time-consuming experiments were avoided.

### 6.3.2 Remarks

A CEKF characterized by a systematic tuning procedure was developed to reconstruct the product composition of a binary distillation column. The obtained results showed that the knowledge of the model errors, in particular of the thermodynamic model, is useful to obtain a reliable estimator avoiding a time-consuming trial and error procedure. The performance of the estimator was tested...
for two different operating conditions and using two different sets of NRTL parameters to describe the vapor liquid equilibrium. The estimated product compositions were successfully compared to experimental data carried out in a pilot plant.
Chapter 7

Conclusions

This Thesis addresses the issue of estimating product composition in distillation columns. This research has been motivated by the limitations in distillation column monitoring due to the difficulties with on-line analyzers. In order to find a robust estimator for distillation column this work started analyzing the main difficulties met in designing estimators for distillation columns. The most important questions investigated were: (i) the choice of a robust estimation structure; (ii) and the estimator tuning. The selection of the estimation structure has been addressed within an adjustable-structure geometric estimation framework. By structure we mean the sensors number and locations as well as of the set of innovated states where the measurement information is directly injected. By means of the sensitivity analysis it was found a passivated structure that jointly combined with NGE or EKF algorithms led to design low order estimators. Soft Sensors with a passivated structure are particularly favorable for columns that work in high purity condition. Indeed, the estimation structure obtained allows to interrupting the mechanism of the error propagation in those column regions where the information mechanism injection is not working, i.e., characterized by a weak relationship between temperature and compositions. This implies the possibility to design an estimator with a modular construction by means the combination of more single-sensor single-innovated scheme distributed along the column. In this way, the information injection mechanism could be connect or disconnected on the basis of the system thermodynamic aspects. Moreover, the design of an adaptive estimation scheme
with the possibility of changing the estimation structure during the column operation was considered. The adaptive-adjustable estimation scheme allowed changing the temperature sensor location during the process separation following the temperature wave. This methodology could be particularly useful for batch column characterized by an intrinsically dynamic behavior.

The approach was tested with an experimental 32-stage pilot plant where the separation of a mixture of ethanol/tert-butanol/water takes place, yielding the following results: (i) one sensor used, (ii) two states are innovated, and (ii) with this structure, non linear geometric (NGE with 66 equations) and the reduced Extended Kalman Filter obtained (GEKF) with 67 equations yield to the same results obtained with a CEKF (with 2144 ODE’s).

Finally, in order to overcome the difficulties related to the time-consuming trial and error tuning procedure an alternative tuning procedure has been proposed. This technique is based on an approximation of the model covariance matrix and allows obtaining a more systematic tuning of the Kalman filter.
Bibliography


Lundstrom, P.; Studies on Robust Multivariable Distillation Control, PhD Thesis University of Trondheim.


Skogestad S.; Postlethwaite I. (1996); Multivariable feedback and control: analysis and design. Wiley.


