



POSGRADO EN CIENCIAS APLICADAS



**TWO SCALES OF BIOCHEMICAL REACTIONS:  
BIOREACTORS AND GENE REGULATION NETWORKS**

TESIS QUE PRESENTA

**Vrani Ibarra Junquera**

PARA OBTENER EL GRADO DE DOCTOR EN CIENCIAS APLICADAS

EN LA OPCIÓN DE CONTROL Y SISTEMAS DINÁMICOS

Codirectores de tesis:

**Dr. Haret Codratian Rosu Barbus**

**Dr. Arturo Zavala Río**

San Luis Potosí, S.L.P., México, julio de 2006



## Constancia de aprobación de la tesis

La tesis "*Two Scales of Biochemical Reactions: Bioreactors and Gene Regulation Networks*" presentada para obtener el Grado de Doctor en Ciencias Aplicadas en la opción de Control y Sistemas Dinámicos fue elaborada por Vrani Ibarra Junquera y aprobada el 6 de julio de 2006 por los suscritos, designados por el Colegio de Profesores de la División de Matemáticas Aplicadas y Sistemas Computacionales del Instituto Potosino de Investigación Científica y Tecnológica, A.C.

---

**Dr. Haret Codratian Rosu Barbus**  
Co-director de la tesis

---

**Dr. Arturo Zavala Río**  
Co-director de la tesis

---

**Dr. José Salome Murguía Ibarra**  
Presidente del jurado

---

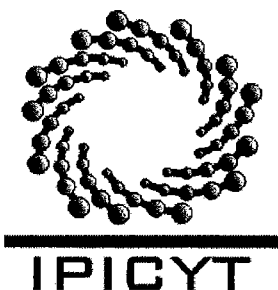
**Dr. Román López Sandoval**  
Secretario del jurado



## **Créditos Institucionales**

Esta tesis fue elaborada en la División de Matemáticas Aplicadas y Sistemas Computacionales del Instituto Potosino de Investigación Científica y Tecnológica, A.C., bajo la dirección del Dr. Haret Codratian Rosu Barbus y del Dr. Arturo Zavala Río.

Durante la realización del trabajo el autor recibió una beca académica del Consejo Nacional de Ciencia y Tecnología (165315) y del Instituto Potosino de Investigación Científica y Tecnológica, A. C.



# Instituto Potosino de Investigación Científica y Tecnológica, A.C.

## Acta de Examen de Grado

COPIA CERTIFICADA

El Secretario Académico del Instituto Potosino de Investigación Científica y Tecnológica, A.C., certifica que en el Acta 002 del Libro Primero de Actas de Exámenes de Grado del Programa de Doctorado en Ciencias Aplicadas en la opción de Control y Sistemas Dinámicos está asentado lo siguiente:

En la ciudad de San Luis Potosí a los 6 días del mes de julio del año 2006, se reunió a las 17:00 horas en las instalaciones del Instituto Potosino de Investigación Científica y Tecnológica, A.C., el Jurado integrado por:

<b>Dr. José Salomé Murguía Ibarra</b>	<b>Presidente</b>	<b>UASLP</b>
<b>Dr. Román López Sandoval</b>	<b>Secretario</b>	<b>IPICYT</b>
<b>Dr. Arturo Zavala Río</b>	<b>Sinodal</b>	<b>IPICYT</b>
<b>Dr. Haret-Codratián Rosu Barbus</b>	<b>Sinodal</b>	<b>IPICYT</b>

a fin de efectuar el examen, que para obtener el Grado de:

**DOCTOR EN CIENCIAS APLICADAS  
EN LA OPCIÓN DE CONTROL Y SISTEMAS DINÁMICOS**

sustentó el C.

**Vrani Ibarra Junquera**

sobre la Tesis intitulada:

*Two Scales of Biochemical Reactions: Bioreactors and Gene Regulation Networks*

que se desarrolló bajo la dirección de


**Dr. Arturo Zavala Río**  
**Dr. Haret-Codratián Rosu Barbus**

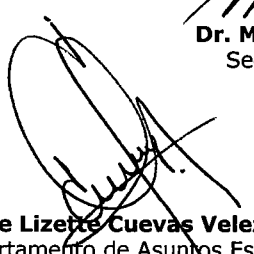
El Jurado, después de deliberar, determinó

**APROBARLO**


Dándose por terminado el acto a las 18:45 horas, procediendo a la firma del Acta los integrantes del Jurado. Dando fé el Secretario Académico del Instituto.

A petición del interesado y para los fines que al mismo convengan, se extiende el presente documento en la ciudad de San Luis Potosí, S.L.P., México, a los 6 días del mes julio de 2006.

  
**Dr. Marcial Borja Mari**  
Secretario Académico



**L.C.C. Ivonne Lizette Cuevas Velez**  
Jefa del Departamento de Asuntos Escolares

  
INSTITUTO POTOSINO  
DE INVESTIGACIÓN  
CIENTÍFICA Y TECNOLÓGICA, A.C.  
**IPICYT**  
SECRETARIA ACADEMICA

---

By definition, biochemical engineers are concerned with biochemical systems, often with systems employing growing cells. Even the simplest living cell is a system of such forbidding complexity that any mathematical description of it is an extremely modest approximation.

**James E. Bailey**

---

## Abstract

In this thesis an attempt is made to show the advantages of applying dynamical analysis and control techniques in the field of modern engineering of biochemical reactions. In this sense, we identify two main problems at two different scales which from a top bottom perspective are the following: the dynamical behavior of bioreactors and the gene regulation processes. With respect to the first subject, we tackle three important issues: i) the dynamical behavior of controlled enzymatic reactors, ii) the following of optimal batch operation models in fed-batch bioreactors, and iii) the possibility to infer mixed culture growth from a total biomass signal. With regard to the bottom subject (gene regulation), the potential usage of a software sensor for monitoring these fundamental processes is developed in detail. It is worth mentioning that the latter development was included in the selected list of contemporary topical areas in biological physics studies of the *Virtual Journal of Biological Physics Research* in August 2005 (Vol. 10, Issue 3).

**Keywords:** Bioreactors, Gene regulation, Dynamical analysis, Control

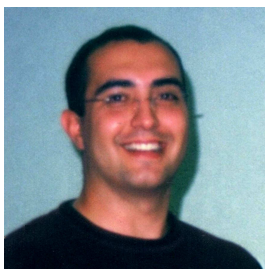
## Resumen

Este documento de tesis pretende mostrar las ventajas del uso del análisis dinámico y las técnicas de control automático en el campo de la moderna ingeniería de las reacciones bioquímicas. Es en este sentido que identificamos dos problemas centrales a dos escalas distintas, los cuales pueden ser clasificados desde una perspectiva de arriba-abajo como: el comportamiento dinámico de bioreactores y el proceso de regulación genética. Con respecto al primer tema, nosotros afrontamos tres cuestiones: i) el comportamiento dinámico de un biorreactor enzimático controlado, ii) el seguimiento de modelos de lote óptimos en biorreactores de tipo lote alimentado, iii) la posibilidad de inferir crecimientos mixtos a partir de una señal de biomasa total. Con respecto al tema de regulación genética, se muestra en detalle el potencial uso de sensores computacionales en el monitoreo de estos procesos fundamentales. Cabe mencionar, que este último trabajo fue escogido para aparecer en la seleccionada lista de publicaciones de áreas de actualidad de *Virtual Journal of Biological Physics Research* en agosto del 2005 (Vol. 10, Issue 3).

**Palabras claves:** Biorreactores, Regulación genética, Análisis dinámico, Control

---

## The Author



**Vrani Ibarra-Junquera** is a Biochemical Engineer [ITM, Mexico, (1995-2000)]. His master degree is in Chemical Engineering, obtained in UASLP, Mexico, (2001-2003). His main research interests are in dynamics and control of biochemical reactors, gene networks and chemotaxis. e-mail: vrani@ipicyt.edu.mx.

## The Supervisors



**Dr. Haret C. Rosu** is a member of the National System of Researchers in Mexico (SNI-II) and in the staff of IPICyT since 2002. His research interests are in mathematical physics and mathematical modeling of biological systems. He obtained his PhD degree in Physics in the Institute of Atomic Physics, campus Magurele-Bucarest, Romania. e-mail: hcr@ipicyt.edu.mx



**Dr. Arturo Zavala Río** is a member of the National System of Researchers in Mexico (SNI-I) and in the staff of IPICyT since 2001. His research interests are in automatic control and dynamical analysis of nonlinear systems. His PhD degree in Automatic Control is from Grenoble National Polytechnic Institute (INPG), France. e-mail: azavala@ipicyt.edu.mx

---

## Acknowledgements

I owe much gratitude to my supervisor, Professor Dr. Haret C. Rosu, for his invaluable guidance and support, but mostly for his unparalleled dedication and enthusiasm. Thanks also go to my co-supervisor, Dr. Arturo Zavala R o, for all the help and support which allowed this academic achievement.

I thank my wife Pili whose unwavering love and support have instilled in me the courage to strive for success. My Mother and Father have always been with me and their love, support, and exemplary lives have encouraged me in all my academic and personal goals.

I would also like to thank CAPEC for a seven-month stay in 2005, during which a part of this thesis has been accomplished. I am also grateful to Professor Dr. Sten Bay J rgensen and Professor Dr. Rafiqul Gani for all their help during that stay.

Last but not least, I thank CONACyT twice: as participant of the project 46980 and especially for the doctoral fellowship that made possible the academic achievements that lead to this thesis.

Vrani Ibarra Junquera



---

---

# Contents

---

<b>1</b>	<b>Preface and General Introduction</b>	<b>1</b>
<b>2</b>	<b>PI-Controlled Bioreactor</b>	<b>5</b>
2.1	Proportional and Integral (PI) Controlled Bioreactor as a Generalized Liénard System . . . . .	5
2.2	Introduction . . . . .	5
2.3	Cholette's Dynamical Model . . . . .	7
2.4	Transformation to the Liénard form . . . . .	8
2.5	Existence of Limit Cycles . . . . .	9
2.6	Uniqueness of Limit Cycles . . . . .	11
2.7	Case Study . . . . .	13
2.8	A Local Center of the Generalized Liénard System ? . . . . .	15
2.9	Conclusions . . . . .	16
<b>3</b>	<b>Optimal Batch Operation Models</b>	<b>18</b>
3.1	Following the Optimal Batch Bioreactor Operations Model . . . . .	18
3.2	Introduction . . . . .	19
3.3	Finding the Optimal Batch Operation Model . . . . .	20
3.3.1	The optimal batch operation model . . . . .	22
3.3.2	Pontryagin's formulation . . . . .	22
3.3.3	Active path constraints . . . . .	23
3.3.4	Solution inside the feasible region . . . . .	23
3.3.5	The case <i>i)</i> and <i>ii)</i> . . . . .	25
3.3.6	The case <i>iii)</i> . . . . .	25
3.4	The Ideal Model . . . . .	27

3.5	The Realistic Model . . . . .	33
3.6	Following the Optimal Batch Operation Model . . . . .	35
3.6.1	Reachability Analysis . . . . .	35
3.6.2	Studying the Exact Feedback Linearization Problem . . . . .	37
3.6.3	Robust Nonlinear Synchronization . . . . .	39
3.6.4	Robust Control Law . . . . .	41
3.7	Concluding Remarks . . . . .	45
3.8	Notation and Definitions . . . . .	46
3.9	Proof of proposition 3.1 . . . . .	47
<b>4</b>	<b>Inferring Mixed Culture Growth</b>	<b>55</b>
4.1	Inferring mixed-culture growth from total biomass data in a wavelet approach . . . . .	55
4.2	Introduction . . . . .	55
4.3	A simple mixed-growth model . . . . .	57
4.4	Mixed cultures on mixtures of substrates . . . . .	59
4.5	Measuring regularity with the wavelet transform . . . . .	60
4.5.1	The wavelet transform . . . . .	60
4.5.2	Wavelet singularity analysis . . . . .	62
4.6	Results and discussion . . . . .	63
4.6.1	Wavelet analysis for the MCMS case . . . . .	64
4.6.2	Wavelet analysis for the noisy data case . . . . .	64
4.7	Concluding remarks . . . . .	65
<b>5</b>	<b>Gene Regulation Networks</b>	<b>77</b>
5.1	Nonlinear Software Sensor for Monitoring Genetic Regulation Processes with Noise and Modeling Errors . . . . .	77
5.2	Introduction . . . . .	78
5.3	Brief on the biological context . . . . .	79
5.4	Mathematical Model for Gene Regulation . . . . .	81
5.5	Nonlinear Software Sensor . . . . .	82
5.6	Three-Gene Circuit Case . . . . .	87
5.7	Conclusion . . . . .	91
<b>I</b>	<b>Bibliography</b>	<b>94</b>

---

## Preface and General Introduction

---

Recently, the concept of industrial biotechnology is strongly emphasized and developed by many scientists, chemical companies, as well as governments. From the technological point of view, this is not a new concept but its recent impetus is mainly due to biocatalysis and fermentation in combination with recent breakthroughs in the forefront of molecular biology and metabolic engineering of industrially useful microorganisms.

Moreover, bioprocesses are ubiquitous – they range from the production of food in a kitchen to the synthesis of sophisticated and extremely valuable therapeutics–. In addition, bioprocesses involve the use of cells (or cellular molecules) as micro-factories to manufacture the product of interest. Natural cellular products can be directly utilized or, alternatively, the cells can be engineered to generate products of interest. In either case, the production levels can be improved by employing several strategies for processing and/or control. However, to get an enhanced profit from those micro-factories the analysis and understanding of the transcriptome and proteome is of extreme importance, because we can infer the underlying reaction mechanisms between genotype and phenotype. Such type of knowledge provides substantial information for controlling the metabolism at the biochemical level. Accurate metabolic models could be the background for monitoring and even controlling indirectly the cellular metabolism.

However, it is fair to mention that the majority of biologists use the term *model* as a formal description of connections and dependencies between different parts of a metabolic pathway. In addition, they commonly use verbal descriptions, instead of mathematical equations, and their models usually are limited to qualitative descriptions of the connections, without including quantitative relationships. On the other hand, for systems theorists, *model* means a system of equations, with initial and/or boundary conditions, that (ideally) describes the qualitative and quantitative behavior of a system, including its dynamics over a full range of values. Therefore,

the fundamental differences between mathematical models of molecular biologists and systems theorists lie in the range of dynamical behavior that they expect to be able to represent by means of their model. In the case of biology, most models still only cover steady-state properties of a system. Nevertheless, recently, the mathematical dynamical models based on ordinary differential equations are a very useful abstraction because they can provide qualitative and quantitative predictions. In this thesis, we use the models and their dynamical analysis to gain insight into the mechanisms underlying the observed phenomena, predict patterns of behavior and design strategies for their automatic control. As most of the real life problems the biological processes involve nonlinear behavior. Thus, the dynamical analysis of nonlinear systems is quite valuable in the field of biochemical engineering, especially from the point of view of design and control. This was the main task in our research.

From our point of view, it is possible to develop a framework using dynamic analysis and control techniques to qualitatively gain insight on the link between the genomic level and the physiology of whole organisms. We are particularly interested in the analysis, monitoring and control of biological dynamical processes. We think that a good procedure is to start with a genomically detailed model of a subsystem of interest. Next, insert this detailed submodel into a cellular model with some pseudo-molecular details, and then employ such a cellular model in the whole process system model.

For instance, it is well known that mixing is an important natural as well as technological process in chemical engineering. This is even more relevant when biochemical reactions get involved. We will refer to the case of continuous stirred tank reactors (CSTRs), for which Lo and Cholette [2] developed a nonideal isothermal mixing model using a Haldane type chemical reaction rate (which is similar to the Monod function for low concentrations but includes the inhibitory effect at high concentrations). In this thesis, we will show that the study of such a model under closed loop condition by means of a proportional and integral controller is relevant. This will be the topic treated in the first chapter.

Another interesting problem is how to follow an optimal batch operations model for a bioreactor in the presence of uncertainties. The optimal batch bioreactor operation model (OBBOM) refers to the bioreactor trajectory that is required to be followed such that the nominal cultivation be optimal. In that sense, this thesis presents an original idea to solve this type of problem. We consider the optimal operations model as the master system which has the optimal cultivation trajectory for the feed flow rate and the substrate concentration. On the other hand, the "real" bioreactor, the one with unknown dynamics and perturbations, is considered as the slave system. Then, we design a control law that ensures the synchronization of both systems.

Other important biochemical processes are microbial growths; their modeling is a problem of special interest in mathematical biology and theoretical ecology, biotechnology and bioengineering. A significant class of such processes is the traditional fermented foods and beverages

in which either endemic microorganisms or inocula with selected microorganisms are used. However, the phenomenological details and the theory of the time evolution of the fermentation are as yet poorly understood. Here, we present for the very first time a wavelet singularity approach to infer such type of growth.

Besides, historically, biologists have tried to understand organisms by investigating progressively smaller details within them in order to gain a better understanding of the general concepts. Recently, there is a trend to look for properties that emerge when groups of such elementary components interact. For instance, the gene expression is a complex dynamic process with intricate regulation networks. Currently, it is well accepted that the expression of genes may be regulated at several levels, from transcription to RNA, and even as post-translational modification of protein activity. We do think that the techniques of nonlinear dynamical analysis have tremendous potential in the investigations focusing on gene expression processes. Unfortunately, the control and in general the dynamic terminology and the whole set of mathematical methods are poorly known to the majority of biochemists. Therefore, one of the aims of this work is to put forth the diffusion and understanding of such mathematical techniques in this area. This is one of the goals of the last chapter of this document.

The thesis is organized as follows. In the next chapter, we tackle the mathematical problem of identifying limit cycles in a continuous bioreactor. Then, the third chapter treats the problem of inducing optimal trajectories in a fed-batch bioreactor. The fourth chapter deals with the problem of identifying mixed growth from total biomass data. Finally, in the fifth chapter, we attempt to go deeper in the biological processes governing the cell cycle, i.e., the gene regulation process.

The present thesis has been structured is base of the following papers:

- **Chapter 2**  
V. Ibarra-Junquera & H.C. Rosu, (2006), *PI-Controlled Bioreactor as a Generalized Lié-nard System*. Accepted in Computers & Chemical Engineering. (arXiv: nlin.CD/0502049 v2).
- **Chapter 3**  
V. Ibarra-Junquera, & S.B. Jørgensen, (2006), *Following the Optimal Batch Bioreactor Operations Model*. To be submitted to Computers & Chemical Engineering.
- **Chapter 4**  
V. Ibarra-Junquera, P. Escalante-Minakata, J.S. Murguía-Ibarra, & H.C. Rosu, (2006), *Inferring mixed-culture growth from total biomass data in a wavelet approach*. Accepted in Physica A. (arXiv.org:physics/0512186).
- **Chapter 5**  
V. Ibarra-Junquera, L.A. Torres, H.C. Rosu, G. Argüello, & J. Collado-Vides, (2005), *Nonlinear software sensor for monitoring genetic regulation processes with noise and modeling errors*. Phys. Rev. E 72, 011919. (arXiv: physics/0410096).

---

---

## PI-Controlled Bioreactor

---

### 2.1 Proportional and Integral (PI) Controlled Bioreactor as a Generalized Liénard System

This work was performed in collaboration with Professor Dr. Haret C. Rosu [1]. Here we show that periodic orbits can emerge in Cholette's bioreactor model working under the influence of a PI-controller. We find a diffeomorphic coordinate transformation that turns this controlled enzymatic reaction system into a generalized Liénard form. Furthermore, we give sufficient conditions for the existence and uniqueness of limit cycles in the new coordinates. We also perform numerical simulations illustrating the possibility of the existence of a local center (period annulus). A result with possible practical applications is that the oscillation frequency is a function of the integral control gain parameter.

### 2.2 Introduction

Mixing, understood as interpenetration of particles in different zones of a given volume, is an important natural as well as technological process. This is even more so when biochemical reactions get involved. For the case of continuous stirred tank reactors (CSTRs), Lo and Cholette [2] developed a nonideal isothermal mixing model using a Haldane type chemical reaction rate (which is similar to the Monod function for low concentrations but includes the inhibitory effect at high concentrations). This model has been studied extensively later by many authors ([3], [4], [5], [6], [7]). In particular, Sree and Chidambaram [5], [6] focused on the control problem by means of a proportional integral (PI) control for this case. Indeed, the PI controller is broadly used in the chemical and biochemical industry. Therefore its closed-loop behavior is of much interest. In this chapter, we present a novel mathematical feature of this

closed-loop enzymatic reaction system, namely the possibility to be represented as a dynamical system corresponding to a non polynomial Liénard oscillator. That means that given a PI-controlled CSTR governed by the usual two-dimensional smooth dynamical system

$$\dot{X} = P(X, Y), \quad \dot{Y} = Q(X, Y), \quad (2.1)$$

we are able to find a diffeomorphic coordinate transformation (Eq. (2.7) below) that allows us to put it into the well-known generalized Liénard form

$$\dot{X} = \phi(Y) - F(X) \quad (2.2)$$

$$\dot{Y} = -g(X), \quad (2.3)$$

where  $g(X)$  is continuous on an open interval  $(a_1, b_1)$ , the functions  $F(X)$  and  $\phi(Y)$  are continuously differentiable on the open intervals  $(a_1, b_1)$  and  $(a_2, b_2)$ , respectively. In fact, these intervals can be extended to  $-\infty < a_i < 0 < b_i < \infty, i = 1, 2$ .

In this chapter, we show that the PI-controlled Cholette's CSTR model belongs to this class of generalized Liénard systems. Once doing this, we make use of the results encountered in this research area to study the periodic solutions near stationary points for this particular application. On the other hand, the Hopf bifurcation is an efficient way to study the existence of periodic orbits. In this case, a pair of complex eigenvalues of the jacobian matrix, evaluated at the unique equilibrium point, exist and to cross transversally the imaginary axis. Nevertheless, the fact that a Hopf bifurcation guarantees the existence of a limit cycle does not imply its uniqueness [8] within the state space. It is here where the uniqueness result for Liénard systems comes into play.

Thus, we extend the area of application of the Liénard-type system to the case of PI-controlled bioreactors for which we present results on the existence of limit cycles and its uniqueness as a function of the control gains. The study of oscillatory behavior in bioreactors is might be important issue since it is generated by the coupled dynamics of the most common controller in industry (the PI one) and the kinetics of the biochemical reactions. In addition, we shed light here on an explicit example of a closed-loop system which is of Liénard-type. Since the most direct way to influence the PI-controlled Cholette's CSTR is through the gain of the controller, the present analysis provide the users with definite conditions for inducing oscillatory behaviors. Such conditions could be instructive from the pedagogical standpoint as well.

The chapter is organized as follows. In Section 2.3, we discuss the PI-controlled Cholette's bioreactor model and its basic assumptions. In Section 2.4, we present the coordinate transformation that leads to the Liénard representation of this type bioreactor. The existence of limit cycles is discussed in Section 2.5, and the uniqueness consideration is included in Section 2.6. The section 2.7 shown the case studied. The numerical simulation that we performed indicating the possible presence of the period annulus is shortly described in Section 2.8. Finally, we end up the chapter with several concluding remarks.



## 2.3 Cholette's Dynamical Model

The dynamical behavior without control actions (i.e., open-loop operation) is governed by a unique nonlinear ordinary differential equation (see Eq. (2.4)) [3]. The nonideal mixing can be described by the Cholette model [3]. This model was studied by Chidambaram [5], who proposed a tuning method for a PI-controller. Examples, where this kind of kinetics occurs, can be found in [7]. The reactor model is given by the following equations

$$\frac{d\zeta}{dt} = (\zeta_f - \zeta) \left[ \frac{nF}{mV} \right] - \frac{K_1 \zeta}{(1 + K_2 \zeta)^2}, \quad (2.4)$$

where the parameters and variables meaning is given Table 2.1.

Table 2.1: Variables and parameters of Cholette's model.

Symbol	Meaning	Units
$\zeta$	Substrate concentration	$[Kmol/m^3]$
$Y$	Integrated error	$[Kmol s/m^3]$
$F$	Feed flow rate	$[m^3/s]$
$V$	Volume	$[m^3]$
$\zeta_F$	Substrate feed concentration	$[Kmol/m^3]$
$K_1$	Maximal kinetic rate	$[1/s]$
$K_2$	Inhibition parameter	$[m^3/Kmol]$
$n$	Mixing parameter	$[dimensionless]$
$m$	Mixing parameter	$[dimensionless]$
$K_C$	Proportional gain controller	$[dimensionless]$
$K_i$	Integral gain controller	$[dimensionless]$
$u$	Control input	$[dimensionless]$

The following assumptions hold in Eq. (2.4): (i) all model parameters and physicochemical properties are constant (ii) the reaction occurs in a nonideal mixed CSTR, operated under isothermal conditions. The fraction of the reactant feed that enters the region of perfect mixing is denoted by  $n$ , whereas  $m$  denotes the fraction of the total volume of the reactor where perfect mixing is achieved. For  $m$  and  $n$  both equal to 1, the system is ideally mixed. The values of the parameters  $m$  and  $n$  can be obtained from the residence time distribution [3]. Fig. 2.1 shows the schematic diagram of the bioreactor configuration modeled by Eq. (2.4).

Following the previous works ([3] [5]), we consider  $\zeta_f$  as the manipulated variable (i.e.  $\zeta_f = u$ ) and let  $\zeta$  be the controlled variable [5]. We are especially interested in the conditions that induce

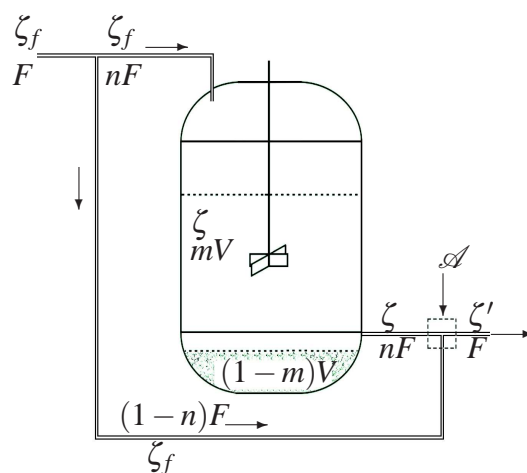


Figure 2.1: Schematics of a classical continuous stirred tank bioreactor with imperfect mixing corresponding to Cholette's model.

oscillatory behavior in the bioreactor. The common control law in this case is of the proportional integral type, that requires an error dynamical extension in order to build the closed-loop system of two dimensions. Thus, the control law is given by

$$u \triangleq \left( -K_c \cdot \text{Error} - K_i \int \text{Error} dt \right),$$

where  $K_c$  and  $K_i$  are the control gain values. For the sake of simplicity, the closed-loop system is written as:

$$\dot{X} = -X \left( C + \frac{K_1}{(1 + K_2 X)^2} \right) + C(-K_c(X - \text{Ref}) - K_i Y) \quad (2.5)$$

$$\dot{Y} = X - \text{Ref}, \quad (2.6)$$

where

$$C = \frac{nF}{mV}, \quad X = \zeta.$$

Eq. (2.5) describes the dynamical behavior of the concentration, while Eq. (2.6) refers to the dynamical behavior of the integrated error.

## 2.4 Transformation to the Liénard form

In this section, we show that the system given by Eqs. (2.5)-(2.6) can be rewritten as a system of the form (2.2)-(2.3), i.e., in the Liénard generalized form. This is one of the main results of this work.

**Proposition 1.** Under the transformation

$$\begin{bmatrix} x \\ y \end{bmatrix} = \Psi(X, Y) = \begin{bmatrix} X - X_p \\ -Y + Y_p \end{bmatrix}, \quad (2.7)$$

where

$$\begin{aligned} Y_p &= -\frac{Ref(C + K_1 + RefCK_2(K_2Ref + 2))}{CK_i(1 + K_2Ref)^2} \\ X_p &= Ref, \end{aligned}$$

system given by Eqs. (2.5)-(2.6) can be written in the generalized Liénard form with the following properties

$$[A1] \quad g(0) = 0 \text{ and } xg(x) > 0 \text{ for } x \neq 0;$$

$$[A2] \quad \phi(0) = 0 \text{ and } \phi'(x) > 0 \text{ for } a_2 < x < b_2;$$

$$[A3] \quad \text{The curve } \phi(y) = F(x) \text{ is well defined for all } x \in (a_1, b_1).$$

**Proof.** If we substitute  $X = x + X_p$  and  $Y = -y + Y_p$  in Eqs. (2.5) and (2.6), and we choose

$$F(x) = x \left( C(1 + K_c) + K_1 \frac{\left( 1 - \frac{K_2Ref(2 + K_2(x + 2Ref))}{(1 + K_2Ref)^2} \right)}{(1 + K_2(x + Ref))^2} \right) \quad (2.8)$$

$$\phi(y) = CK_i y \quad (2.9)$$

$$g(x) = x, \quad (2.10)$$

we get the generalized Liénard form of the PI-controlled Cholette system. The properties [A1], [A2] and [A3] are straightforwardly checked in Eqs. (2.8)-(2.10).  $\square$

For  $\Psi(X, Y)$  to be a local diffeomorphism on the region  $\Omega$ , it is necessary and sufficient that the Jacobian  $d\Psi(X, Y)$  be nonsingular on  $\Omega$ . Since  $\Psi(X, Y)$  is linear and, in this case, the determinant of the Jacobian matrix is constant, then is nonsingular in the region  $\Omega = [-\infty, \infty] \times [-\infty, \infty]$ .

In the literature, the properties [Ai] are standard properties assumed for Liénard systems [11]. We point out that the huge existing literature on Liénard systems deals mainly with cases in which  $F(x)$  is polynomial [15], [13], [14], whereas we are in a case in which  $F(x)$  is a nonlinear rational function. Such cases are far less studied and there are still many open problems.

## 2.5 Existence of Limit Cycles

We briefly recall some basic results of the theory of bifurcations of vector fields. Roughly speaking, a bifurcation is a change in equilibrium points, periodic orbits, or in their stability

properties, when varying a parameter known as bifurcation parameter. The values of the parameter at which the changes occur are called bifurcation points. A Hopf bifurcation is characterized by a pair of complex conjugate eigenvalues crossing the imaginary axis. Now, suppose that the dynamical system  $\dot{X} = f(X, \mu)$  with  $X \in \mathbb{R}^n$  and  $\mu \in \mathbb{R}$  has an equilibrium point at  $X_{eq}$ , for some  $\mu = \mu^H$ ; that is  $f(X_{eq}, \mu^H) = 0$ . Let  $A(\mu) = \frac{\partial f(X^H, \mu)}{\partial X}$  be the Jacobian matrix of the system at the equilibrium point.

Assume that  $A(\mu^H)$  has as single pair of purely imaginary eigenvalues  $S(\mu^H) = \pm i\omega_H$  with  $\omega_H > 0$  and that these eigenvalues are the only ones with the properties  $\Re(S) = 0$ . If the following condition is fulfilled  $\left. \frac{d\Re(S(\mu))}{d\mu} \right|_{\mu=\mu^H} \neq 0$  the Hopf bifurcation theorem states that at least one limit cycle is generated at  $(X_{eq}, \mu^H)$  (see [9]). The condition (2.5) is known as the transversality hypothesis. Considering now the  $n$ th degree characteristic polynomial  $\lambda(S) = S^n + a_1 S^{n-1} + a_2 S^{n-2} + \dots + a_n$ , where all the real  $a_i$  coefficients are positive allows the construction of a Hurwitz matrix  $\mathcal{H}_{n \times n}$ . Then one has the basic result that the characteristic polynomial is stable if and only if the leading principal minors of  $\mathcal{H}_{n \times n}$  are all positive [10].

To search the Hopf bifurcation we have to calculate the equilibrium points of the system (2.2)-(2.3) making equal to zero the right-hand-side of the equation, and taking as bifurcation parameters the values of the PI-control gains. Then, finding the solution with respect to the state vector  $x$  we notice that the closed-loop system has a unique equilibrium point located at the origin.

**Proposition 2.** If the parameter  $K_c$  is such that

$$K_c^H - 2 \frac{\sqrt{CK_i}}{C} < K_c < K_c^H,$$

then the Liénard system (2.2) and (2.3) with the functions given in Eqs. (2.8)-(2.10) has at least one limit cycle. The upper limit  $K_c^H$  is defined in Eq. (2.15).

**Proof.** If we use the Hurwitz criterion to guarantee that the unique equilibrium point is unstable, we evaluate the Jacobian of the system at the origin

$$\mathcal{J}(0, K_c) = \begin{bmatrix} -C - \frac{K_1}{(1+K_2 Ref)^2} + 2 \frac{Ref K_2 K_1}{(1+K_2 Ref)^3} - CK_c & CK_i \\ -1 & 0 \end{bmatrix}. \quad (2.11)$$

Then, from the determinant  $IS - \mathcal{J}(0, K_c)$ , we get the characteristic polynomial  $\lambda(S) = S^2 + a_1 S + a_2$ , where

$$a_1 = C(1 + K_c) - \frac{K_1(K_2 Ref - 1)}{(1 + K_2 Ref)^3} \quad (2.12)$$

$$a_2 = CK_i. \quad (2.13)$$

Then, the Hurwitz matrix is given by

$$\mathcal{H} = \begin{bmatrix} a_1 & 0 \\ 1 & a_2 \end{bmatrix}, \quad (2.14)$$

and its principal minors are  $\mathcal{H}_1(K_c) = a_1$  and  $\mathcal{H}_2(K_c) = a_1 a_2$ . From this formulas, we can note that the stability of the unique equilibrium point depends on the sign of Eq. (2.12). Since all the parameters in the Eqs. (2.12) and (2.13) are positive, we can induce the local stability as a function of the values of the controller gains involved in these equations and then we obtain the bifurcation point as the trivial solution of Eq. (2.12) for  $K_c$  and restricting  $K_i > 0$

$$K_c^H = -1 + \frac{K_1(K_2 \text{Ref} - 1)}{C(1 + K_2 \text{Ref})^3}. \quad (2.15)$$

In order to test the transversality condition, the behavior of the eigenvalues of  $\mathcal{J}(0, K_c)$  in the neighborhood of  $K_c^H$  should be analyzed. Thus, we take  $K_c$  as  $K_c^H + \varepsilon$ , with  $\varepsilon \in \mathbb{R}$ . The transversality condition will be fulfilled if the sign of the equations  $\mathcal{H}_1$  and  $\mathcal{H}_2$  changes when the sign of  $\varepsilon$  changes. Substitution of Eq. (2.15) in the principal minor expressions gives  $\mathcal{H}_1(K_c^H) = \varepsilon(C)$  and  $\mathcal{H}_2(K_c^H) = \varepsilon(K_i C^2)$ .

From the above equations we can appreciate that if we want to have a positive real part of the eigenvalues of the Jacobian matrix (2.11), we need that  $\varepsilon$  be negative. In other words, it is below the value given by Eq. (2.15) where the limit cycles are generated. Using  $K_c^H$ , the eigenvalues of the matrix (2.11) are given by the roots of the characteristic polynomial  $\lambda(S)$ , which are

$$S_1 = -\frac{\varepsilon C}{2} + \frac{\sqrt{\varepsilon^2 C^2 - 4CK_i}}{2} \quad (2.16)$$

$$S_2 = -\frac{\varepsilon C}{2} - \frac{\sqrt{\varepsilon^2 C^2 - 4CK_i}}{2}, \quad (2.17)$$

where we can notice that the eigenvalues are complex with positive real parts, when  $0 > \varepsilon > -2\frac{\sqrt{CK_i}}{C}$  and  $K_i > 0$ .  $\square$

Eq. (2.15) is of main importance, because it corresponds to the Hopf bifurcation and therefore lies in a neighborhood of the value of the parameter where at least one limit cycle is generated. Note, that the Hopf bifurcation by itself can not guarantee the uniqueness of the limit cycle, because more than one limits cycle could appear [8]. Then we need to use additional constraints in order to find the condition for uniqueness.

## 2.6 Uniqueness of Limit Cycles

Xiao and Zhang [11] gave an interesting theorem on the uniqueness of limit cycles for generalized Liénard systems, under the conditions [A1], [A2] and [A3] that allows us to prove a novel property of PI-controlled Cholette bioreactors.

**Theorem 1.** Using the notations  $G(x) = \int_0^x g(x)dx$  and  $f(x) = F'(x)$ , suppose that the system (2.2) - (2.3) satisfies the following conditions:

- (i) there exist  $x_1$  and  $x_2$ ,  $a_1 < x_2 < 0 < x_1 < b_1$  such that  $F(x_1) = F(0) = 0$ ,  $F(x_2) > 0$  and  $G(x_2) \leq G(x_1)$ ;  $xF(x) \leq 0$  for  $x_2 \leq x \leq x_1$ ,  $F'(x) > 0$  for  $a_1 < x < x_2$  or  $x_1 < x < b_1$ , and  $F(x) \neq 0$  for  $0 < |x| \ll 1$ .

(ii)  $F(x)f(x)/g(x)$  is nondecreasing for  $x_1 < x < b_1$ .

(iii)  $\phi'(y)$  is nonincreasing as  $|y|$  increases.

Then the system given by Eqs. (2.2)-(2.3) has at most one limit cycle, and it is stable if it exists. ■ Note that the above result does not guarantee the existence of limit cycles by itself.

**Proposition 3.** System (2.2) and (2.3) with the functions given by the Eqs. (2.8)-(2.10) and with  $Ref > 2/K_2$  and  $K_c = K_c^H + \varepsilon^*$ , where

$$\varepsilon^* = -\frac{K_1(-2 + K_2 Ref)}{2C(1 + K_2 Ref)^3}, \quad (2.18)$$

has a unique limit cycle.

**Proof.** The existence is given by the Proposition 2, and the uniqueness will be proved by showing that the system fulfills the conditions of Theorem 1. Let us take  $Ref^* = 2\kappa/K_2$  with  $\kappa > 1$  which agrees with the condition stated in Proposition 3. Then

$$K_c^H(\varepsilon^*, Ref^*) = -1 + \frac{K_1 \kappa}{C(1 + 2\kappa)^3}.$$

For these values of the parameters, the function  $F(x)$  is simplified to the form

$$F^*(x) = \frac{xK_1(3\kappa + \kappa K_2^2 x^2 - 4\kappa^3 + 1)}{(1 + 2\kappa)^3(1 + K_2 x + 2\kappa)^2}. \quad (2.19)$$

The real roots of  $F^*(x)$  are 0,  $x_1$ , and  $-x_1$ , where

$$x_1 = \frac{\sqrt{\kappa(-1 + \kappa)}(1 + 2\kappa)}{\kappa K_2},$$

which fulfill the property  $F(x_1) = F(0) = 0$ . Moreover, taking  $x_2 = -\xi x_1$ , where  $0 < \xi < 1$ , one can see that  $x_2 < 0 < x_1$ . To evaluate  $F^*(x)$  in the intervals  $x_2 < x < 0$  and  $0 < x < x_1$ , it is sufficient to substitute in Eq. (2.19)  $x = \beta x_i$ , where  $i = 1, 2$  and  $0 < \beta < 1$ . Then we can easily verify that

$$0 > F^*(\beta x_1) = \frac{(-1 + \kappa)(\beta^2 - 1)\kappa K_1 \beta \sqrt{\kappa(-1 + \kappa)}}{(1 + 2\kappa)^2(\kappa + \beta \sqrt{\kappa(-1 + \kappa)})^2 K_2} \quad (2.20)$$

$$0 < F^*(\beta x_2) = -\xi \frac{(-1 + \kappa)(\beta^2 - 1)\kappa K_1 \beta \sqrt{\kappa(-1 + \kappa)}}{(1 + 2\kappa)^2(-\kappa + \beta \sqrt{\kappa(-1 + \kappa)})^2 K_2}. \quad (2.21)$$

Consequently,  $x F^*(x) \leq 0$  for  $x_2 \leq x \leq x_1$ . Moreover,  $G(x) = x^2/2$  and then  $G(x_2) \leq G(x_1)$ . Besides, for  $\theta > 1$  we have

$$0 < F^{*'}(\theta x_1) = \frac{(-1 + \kappa)\left((-1 + 3\theta^2)\kappa + (\theta + \theta^3)\sqrt{\kappa(-1 + \kappa)}\right) K_1 \kappa^2}{(1 + 2\kappa)^3\left(\kappa + \theta \sqrt{\kappa(-1 + \kappa)}\right)^3}.$$

In other words,  $F^*(x) > 0$  for  $x_1 < x < b_1$ . From Eq. (2.8) we can see that  $F^*(x) \neq 0$  for  $0 < |x| \ll 1$ . In addition, since  $\phi'(y) = CK_i$ , no matter how  $|y|$  increases,  $\phi'(y)$  remains constant. Finally, we can write  $\Psi(x) \equiv F^*(x)f(x)/g(x) = (\mathcal{A} + \mathcal{B})/\mathcal{C}$ , where

$$\mathcal{A} = 2 + \kappa (16 \kappa^2 + K_2^3 x^3) - \kappa (1 + 2 \kappa) (x K_2 (3 K_2 x + 8 \kappa + 4) - 12) \quad (2.22)$$

$$\mathcal{B} = 2 \ln(1 + K_2 x + 2 \kappa) (1 + 2 \kappa)^3 (1 + K_2 x + 2 \kappa) \quad (2.23)$$

$$\mathcal{C} = \frac{K_1^2}{2(1 + 2 \kappa)^6 K_2^2 (1 + K_2 x + 2 \kappa)^3}. \quad (2.24)$$

Performing the derivative of  $\Psi(x)$ , evaluating it at  $x = x_1 \alpha$  for  $\alpha > 1$ , and plotting  $\vartheta \Psi'(x_1 \alpha)$ , where

$$\vartheta = \frac{(1 + 2 \kappa)^5 \left( \kappa + \alpha \sqrt{\kappa(-1 + \kappa)} \right)^4 K_2}{K_1^2 \kappa^3}, \quad (2.25)$$

we get the positive function displayed in Fig. 2.2. Thus,  $F^*(x)f(x)/g(x)$  is nondecreasing for  $x_1 < x < b_1$ . In this way, we checked each of the three conditions of Theorem 1. The unique limit cycle can be seen in Fig. 2.3 that illustrates the numerical simulation corresponding to the results obtained until now.

## 2.7 Case Study

With the goal of illustrating the results given in the previous sections, we perform numerical simulation using the values of the parameters given by Chidambaram [6]. All our calculations are performed for a flow characterized by the value of the Damköhler number ( $Da = K_1 V / F$ ) equal to 300, as reported by Sree and Chidambaram [5].

Table 2.2: The values of the parameters of Cholette's model [5].

Symbol	Value	Units
$F$	$3.333 \times 10^{-5}$	$[m^3/s]$
$V$	$10^{-3}$	$[m^3]$
$K_1$	10	$[1/s]$
$K_2$	10	$[m^3/Kmol]$
$n$	0.75	$[dimensionless]$
$m$	0.75	$[dimensionless]$

In Fig. 2.4 a plot of the evolution in time of the variable  $x$  is given for two values of the integral parameter  $K_i$ . It indicates that the oscillation frequency can be manipulated through this control parameter.

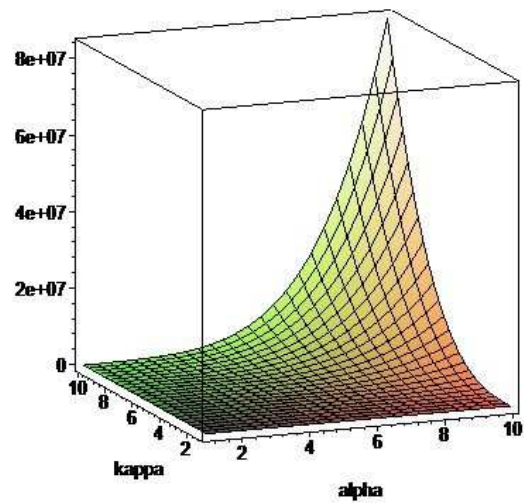


Figure 2.2: The function  $\vartheta\Psi'(\alpha x_1)$ , for  $\alpha > 1$ , and  $\kappa > 1$ . We see that this is a strictly positive function.

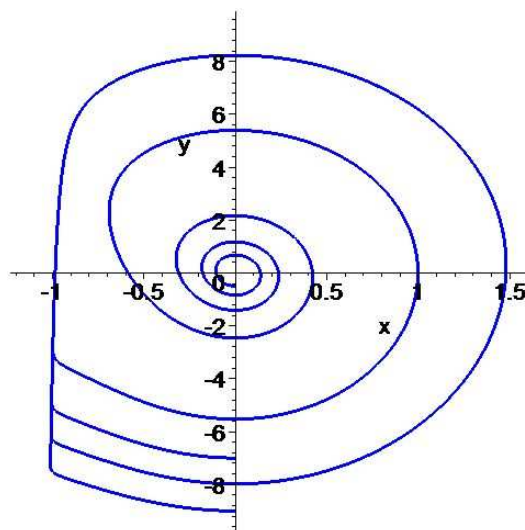


Figure 2.3: The phase portrait of the PI-controlled Cholette model subjected to the uniqueness conditions. The employed values of the parameters are those given in Table 2.2.



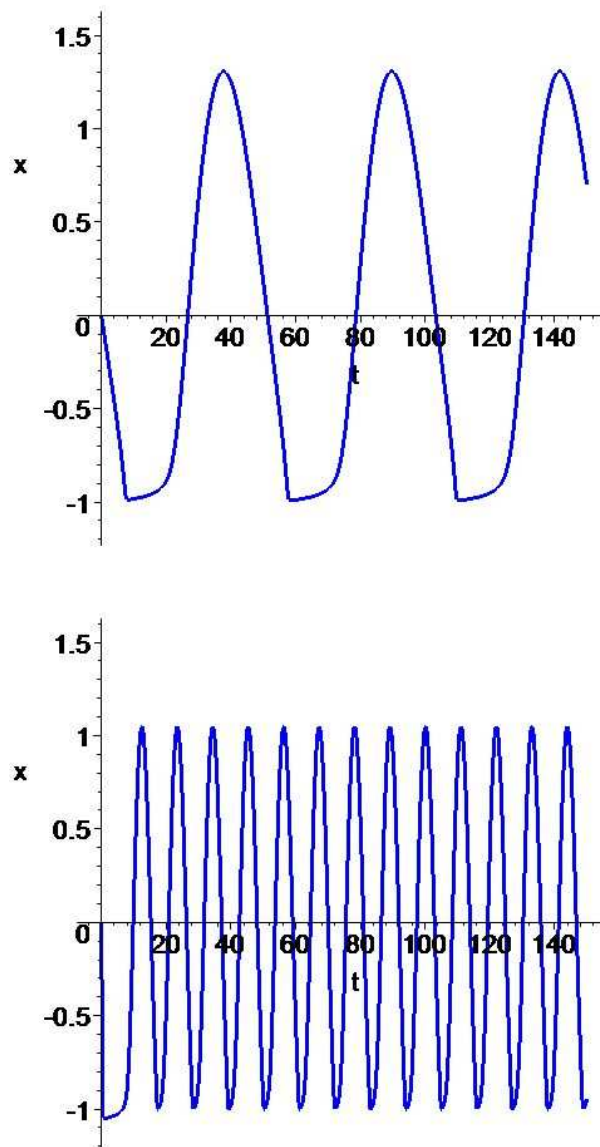


Figure 2.4: Time evolution of the variable  $x$ . The upper plot corresponds to  $K_i = 0.5$ , while the lower plot corresponds to  $K_i = 10$ . This is a graphical representation of the fact that oscillation frequency is a function of the control gain parameter  $K_i$ .

## 2.8 A Local Center of the Generalized Liénard System ?

We begin this section by recalling that a *limit cycle* is an isolated closed orbit, while a critical point is a *center* if all orbits in its neighborhood are closed. To the best of our knowledge the

literature on period annuli for Liénard systems is well developed only for polynomial cases and moreover it focuses on Hamiltonian type systems [12], [16], [17].

Notice that for  $K_c(Ref^*) = K_c^H(Ref^*) + \varepsilon$  and  $\varepsilon$  in the following interval  $\varepsilon^* < \varepsilon < 0$  the existence of limit cycles is proved but without guaranteeing uniqueness. It is precisely in this interval where our numerical simulations point to the existence of a local center in a neighborhood of the origin. Fig. 2.5 shows the phase portrait of the PI-controlled generalized Liénard system with a value of  $K_c$  in the same interval and close to  $K_c^H$ .

## 2.9 Conclusions

The main result of this chapter is that the Cholette CSTR model under PI control can be mapped into a generalized Liénard dynamical system of nonpolynomial type. Thus, we establish a new important application of this class of nonlinear oscillators that allows us to make a detailed study of the oscillatory dynamical behavior of these interesting bioreactors.

Sufficient conditions for the existence and uniqueness of limit cycles of this generalized Liénard system are stated in this chapter together with numerical simulations that indicate the possibility of the existence of a local center (period annulus) when the gain proportional parameter  $K_c$  of the control law is close to the value  $K_c^H$  corresponding to the existence condition of limit cycles. We also notice that the oscillation frequency is a function of the integral control gain parameter  $K_i$ , a result that could have practical applications. We mention that similar results have been obtained by Albarakati, Lloyd, & Pearson [18] for the polynomial case.

Our work also shows that the Liénard representation of dynamical systems and its associated results could have a remarkable potential as an effective tool in the control theory for the closed-loop dynamical analysis in the plane.

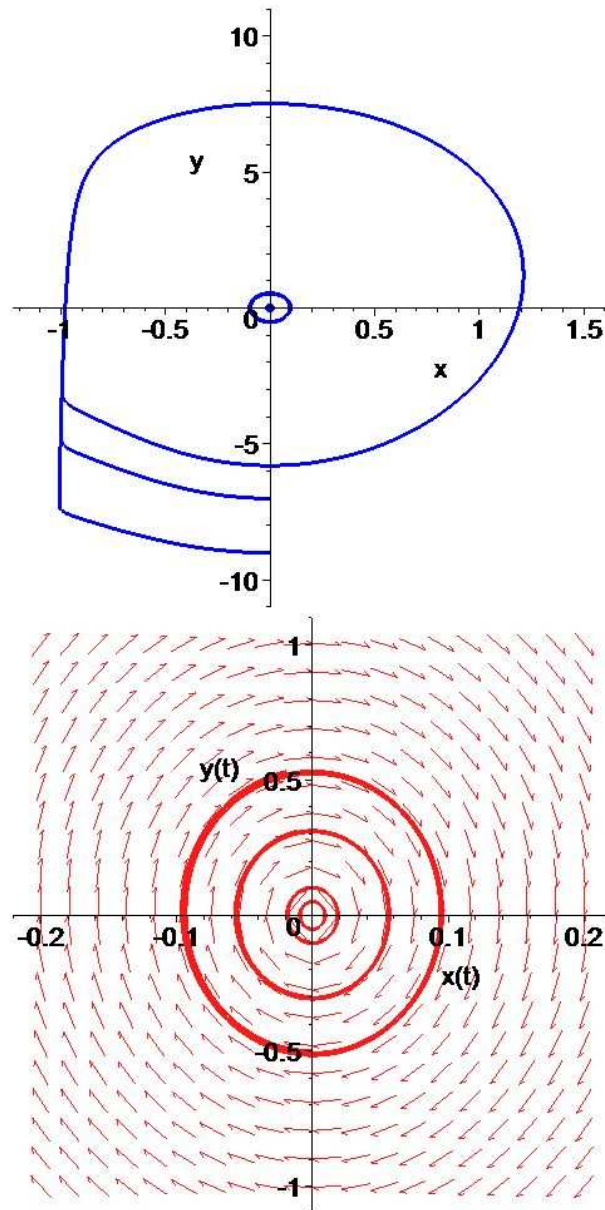


Figure 2.5: The phase portrait of the PI-controlled generalized Liénard system with a value of  $K_c^H(Ref^*) + \varepsilon$  in the interval  $\varepsilon^* < \varepsilon < 0$  and close to  $K_c^H$ . The upper plot shows the state space with the limit cycle and the possible annulus in the neighborhood of the origin, while the lower plot displays the configuration of the vector field close to the origin.

---

---

## Optimal Batch Operation Models

---

### 3.1 Following the Optimal Batch Bioreactor Operations Model

This work was performed in collaboration with Professor Dr. Sten Bay Jørgensen. Here, we study the problem of following an optimal batch operation model for a bioreactor in the presence of uncertainties. The optimal batch bioreactor operation model (OBBOM) refers to the bioreactor trajectory for the nominal cultivation to be optimal. A multiple-variable dynamic optimization of fed-batch reactor for biomass production is studied using a differential geometry approach. The performance index is optimized at the final operation time. The maximization problem for the production of cell mass is solved by finding the optimal filling policy and the optimal substrate concentration in the inlet stream.

In order to follow the OBBOM a master-slave synchronization is used. The OBBOM is considered as the master system which has the optimal cultivation trajectory for the feed flow rate and the substrate concentration. On the other hand, the “real” bioreactor, the one with unknown dynamics and perturbations, is considered as the slave system. The advantage of this idea is that the optimization can easily be performed for a simplified model. Finally, the controller is designed such that the real bioreactor is synchronized with the optimized one in spite of bounded unknown dynamics and perturbations.

As an example, the scheme is applied to a nonlinear fed-batch fermentation process. In addition, this work shows the advantages of using two actuators for control purposes. It is formally proven that the inclusion of the extra input is the solution to the controllability problems pointed out by Szerderkényi et al. [39].

## 3.2 Introduction

From a process systems point of view, the key feature that differentiates continuous processes from batch and semi-batch processes is that the first have a steady state, whereas the latter ones do not. In semi-batch operations, a reactant may be added with no product removal, or the product may be removed with no reactant addition, or a periodic combination of both. Due to the batch dynamics, the process variables need to be adjusted over time to maximize the amount of products and/or quality of interest at the end of the fed-batch cycle. Consequently, this step involves the difficult task of determining time-varying profiles through dynamic optimization.

In general terms, semi-batch processes typically produce low-volume-high-cost products, thus optimal operation is extremely important. Every small improvement in the process may result in considerable reduction in production costs. In addition, it is well known that before a continuous reactor reaches steady state operation, there is a transient period in which it is filled up to the working volume. This period is called the start-up of the reactor and refers to the sequence of operations that leads it from an initial state to a state of continuous operation. The modes of operation usually used for the start-up are sequences of batch and fed-batch modes. Therefore, determining an efficient start-up strategy is important as it saves time. Biochemical reactions are usually slow and an inappropriate start-up may cause the reactor to waste time to reach the desired steady state conditions.

For many chemical and biochemical processes, it is difficult to develop reasonably accurate mathematical models with reliably estimated parameters. This is essential for optimization and design of a high-performance control systems. Especially in biochemical reactors uncertainties are due to the poor process knowledge, nonlinearities, unmodeled dynamics, unknown internal and external noises, environmental influences and time-varying parameters. The presence of such uncertainties causes a mismatch between the formulated optimized model and the true process, which may degrade the performance and can lead to serious control stability problems, especially when the process is nonlinear. Therefore, it is a challenge of significant importance for control engineers to design robust controllers for nonlinear biochemical processes subject to model uncertainties.

Thus, the problem is to design a controller that forces the process measurements to follow an optimal batch operation model in spite of uncertainties. For this purpose a master-slave synchronization scheme is used. That is, the optimal operations model is considered as the master system which provides the trajectory that the feed flow and the substrate concentration should follow in order to achieve an optimal cultivation. On the other hand, the "real" bioreactor (the one with unknown dynamics and perturbations) is considered as the slave. The advantages of this idea are that the optimization can easily be performed for a simplified model and the controller can be designed such that the real bioreactor is synchronized with the master in spite of

bounded unknown dynamics and perturbations. Moreover, sometimes it is not possible to obtain an explicit solution even for simplified models and therefore the synchronization approach allows the possibility to design robust controllers even for those cases.

In order to study this idea, a relatively simple mathematical model is employed. The system consists of a mixing tank, where two streams are fed, one with nutrients at high concentration and the other with a solvent including supplementary components. The reactor is assumed to be perfectly stirred. An unstructured biomass growth rate equation with substrate inhibition kinetics is chosen. Moreover, this work attempts to show the advantages using two actuators for control purposes of the fed-batch cultivations. The structure mentioned above allows the study also of this control problem.

The chapter is organized as follows. In Section 3.3, we discuss the methodology to find the Optimal Batch Operation Model and its basic assumptions. In Section 3.4, the OBBOM for simplified fed-batch bioreactor is founded. In Section 3.5, we present a realistic model of a fed-batch process. An analysis of the realistic system and the methodology for robust track of the OBBOM is discussed in Section 3.6. The concluding remarks are in Section 3.7. Finally, in the appendix we give a few basic preliminaries regarding differential geometry and also the proof of the main statement is addressed.

### 3.3 Finding the Optimal Batch Operation Model

Since [19], the classical linear quadratic regulator techniques in optimal control theory have been successfully applied to linear problems. However, for nonlinear systems some problems still unsolved. This situation is certainly true also for nonlinear control of batch processes.

An interesting work in which the exact solution for a batch bioreactor is derived appears in [20] in which the objective was to maximize the cell productivity. Using the Pontryagin maximum principle for the constant yield case an explicit substrate feeding policy was found for a single feed flow rate case.

In the work [21] a method has been proposed to get an optimal nonlinear control law. They proposed the elimination of the adjoint vector from the repeated time derivatives of the Hamiltonian involving for the first time the Lie brackets of vector fields. The advantage of this method lies in the introduction of a differential operator, a fact that allows obtaining an implicit analytic feedback law expression. Later, in the work [22], the optimal state feedbacks has been characterized in the singular operation region in an analytical manner using tools of differential geometry. The state feedback laws are obtained as a differential operator for invariant-time systems and also extended to time-varying systems. Additionally, the notion of degree of singularity was used, allowing a more transparent characterization of the necessary condition for optimality for the

affine single input case.

Another nonlinear approach was introduced in [23] and later in [26], where an optimal adaptive scheme for a fed-batch fermentation processes involving multiple substrates was investigated. The optimization was achieved using the classical Pontryagin's principle in which the optimal input is obtained as a function of the adjoint variables.

In [24] the state feedback law for on-line optimization problem for single input non-affine systems. Next in [25], it was shown that with a neural network is capable to calculate the switching times for affine single input cases. In [27], the optimal nonlinear input is derived in terms of the states of the system and is independent of the adjoint states. The necessary conditions for optimality are given in terms of the systems Lie brackets and the adjoint states. In addition, the main result is presented in terms of a differential operator.

Using this differential operator in [28] the 2-inputs case was studied. where a nonlinear state feedback laws for on-line optimization of batch processes with one affine and one non-affine input has been introduce. Later, in [30] the multiple-variable dynamic optimization was studied. In the latter work only one affine input and multiple non-affine inputs has been considered. It is stated that the multiple affine inputs case could be treated by transforming the additional affine inputs into non-affine inputs.

In [31], the fed-batch operation of a biochemical reactor is analyzed for a single input case. The feed rate of the substrate is chosen as the control variable. The entire duration of the operation is divided into different subintervals. The system was optimized by approximating the feed flow rate using discrete pulses and a constant flow rate over different subintervals. In the first strategy the equations, lend themselves to an analytical solution. For the second case a shooting method coupled with a sequential quadratic programming technique to obtain the constant flow rates in the different subintervals was used. The method was analyzed for both equal and unequal duration of subintervals.

Another interesting approach appears in [33] for the single input case in which the rate of biomass production was optimized for a predefined feed exhaustion using the residue ratio as a degree of freedom. The analytical expressions of these transitions for variable bioreaction kinetic parameters were determined. In fed-batch operation, the proposed constant feed policy approximates the optimal feed policy closely.

Recently, in [35], the first paper of a two part work, the structure of the optimal solution of a multi-input batch process is characterized in the absence of modeling error. While in the second part, [36], the role of measurements is analyzed to handle uncertainty in the computation of the optimal solution of a multi-input batch process.

An overview of optimal adaptive control of chemical and biochemical reactors is given by Smets et al. [37]. Following the Minimum Principle of Pontryagin the derivation of optimal control sequences for fed-batch production processes is revisited. Extensions towards fermentation processes with multiple substrates and non-monotonic kinetics are also included. In analogy to the work of Vam Impe and Bastin [26] the input is obtained as a function of the adjoint variables. However despite these attempts a robust solution to following the optimal batch operation model is still not available. Therefore this problem is addressed in this chapter.

Taking into account the ever increasing industrial competition, process optimization provides a unified framework for reducing production cost, meeting safety requirements and environmental regulations and also improving product quality, and reducing product variability.

### 3.3.1 The optimal batch operation model

The objective is to optimize a function at the end of the semi-batch cycle. Such problems are called end-point optimization problems and for multi-input affine nonlinear systems can be formulated as follows:

Find an admissible control which minimize the performance index:

$$J = \phi(X(t_f)) \quad (3.1)$$

subject to the dynamics

$$\dot{X} = \mathcal{F}(X, u) = f(X) + \sum_{i=1}^m g_i(X)u_i \quad (3.2)$$

and

$$\mathbf{S}(X, u) \leq 0, \quad (3.3)$$

where  $X \in \mathbb{R}^n$ ,  $\mathbf{S}$  is a  $\zeta$ -dimensional vector of path constraints and  $f$  and  $g_i$  are smooth vector functions,  $u \in \mathbb{R}^m$ ,  $\phi$  is a smooth scalar function representing the terminal cost. In others words, the admissible control should enable the given system to follow an admissible trajectory which at the same time minimizes the performance index.

### 3.3.2 Pontryagin's formulation

By Pontryagin's principle, the problem formulation (3.1)-(3.2) is equivalent to minimizing the Hamiltonian:

$$H = \langle \lambda, \mathcal{F} \rangle + \langle \gamma, \mathbf{S} \rangle = \lambda^T \left( f(X) + \sum_{i=1}^m g_i(X)u_i \right) + \gamma^T \mathbf{S}(X, u), \quad (3.4)$$



where  $\lambda \in \mathbb{R}^n$  is the vector of adjoint variables defined by

$$\dot{\lambda} = -\frac{\partial H}{\partial X}; \quad \lambda(t_f) = \frac{\partial \phi(X)}{\partial X} \Big|_{t_f}; \quad \gamma^T \mathbf{S} = 0. \quad (3.5)$$

A necessary condition for minimizing the Hamiltonian is:

$$\frac{\partial H}{\partial u_i} = \lambda^T \frac{\partial \mathcal{F}}{\partial u_i} + \gamma^T \frac{\partial \mathbf{S}}{\partial u_i} = 0, \quad \forall t \in (0, t_f), \quad \forall i \in (1 \dots m). \quad (3.6)$$

It is said that the solution to the optimization problem is singular in the cases where the maximum principle does not lead to a well-defined relation between the state and the control variable. From the Eq. (3.4), we come to the conclusion that

$$\frac{\partial H}{\partial u_i} = \lambda^T g_i(X) + \gamma^T \frac{\partial \mathbf{S}}{\partial u_i} = 0, \quad \forall i \in (1 \dots m). \quad (3.7)$$

### 3.3.3 Active path constraints

When the input  $u_i$  is computed from an active path constraints, this part of the optimal solution does not depend on the adjoint variables. Each path constraints  $\mathbf{S}_j(X, u)$  is differentiated along the trajectories of Eq. (3.2)

$$\begin{aligned} \frac{\mathbf{d}}{\mathbf{d}t} (\mathbf{S}_j(X, u)) &= L_f \mathbf{S}_j(X, u) + \sum_{i=1}^m L_{g_i} \mathbf{S}_j(X, u) u_i + \sum_{i=1}^m \frac{\partial}{\partial u_i} (\mathbf{S}_j(X, u)) \dot{u}_i \\ \frac{\mathbf{d}^k}{\mathbf{d}t^k} (\mathbf{S}_j(X, u)) &= L_f \left( \frac{\mathbf{d}^{k-1}}{\mathbf{d}t^{k-1}} (\mathbf{S}_j(X, u)) \right) + \sum_{i=1}^m L_{g_i} \left( \frac{\mathbf{d}^{k-1}}{\mathbf{d}t^{k-1}} (\mathbf{S}_j(X, u)) \right) u_i \\ &\quad + \sum_{i=1}^m \sum_{\alpha=0}^n \frac{\partial}{\partial u_i^{(\alpha)}} \left( \frac{\mathbf{d}^{k-1}}{\mathbf{d}t^{k-1}} (\mathbf{S}_j(X, u)) \right) u_i^{(\alpha+1)}, \end{aligned}$$

where  $k \in N$  and the time differentiation of  $\mathbf{S}_j(X, u)$  is continued until the input  $u_i$  appears explicitly. And this  $u_i$  obtained from  $\frac{\mathbf{d}^k}{\mathbf{d}t^k} (\mathbf{S}_j(X, u)) = 0$  represents the optimal input in zone of active path constraints. If  $k \rightarrow \infty$ , then  $u_i$  does not influence the constraint  $\mathbf{S}_j$ , and, thus  $u_i$  can not be obtained from  $\mathbf{S}_j$ .

### 3.3.4 Solution inside the feasible region

When the optimal solution is inside the feasible region, i.e. no constraints are active, the optimal solutions can be obtained from

$$\lambda^T \frac{\partial \mathcal{F}}{\partial u_i} = 0, \quad \forall t \in (0, t_f), \quad \forall i \in (1 \dots m) \quad (3.8)$$

whose time derivatives lead to an infinite number of necessary conditions, that is

$$\frac{\mathbf{d}^k}{\mathbf{d}t^k} \left( \frac{\partial H}{\partial u_i} \right) = \frac{\mathbf{d}^k}{\mathbf{d}t^k} (\lambda^T g_i(X)) = 0, \quad \forall i \in (1 \dots m). \quad (3.9)$$

Now, an operation between vector fields is introduced with the aim of building an algorithmic solution to the optimization problem.

**Definition:** This operation involves two vector fields, the dynamics of the system,  $\mathcal{F} = f + \sum_{i=1}^m g_i u_i$  and  $g_j$ , both defined in a subset  $U \in \mathbb{R}^n$ . From these, a new smooth vector field is constructed, denoted by  $Ad(\mathcal{F}, g_j)(X)$  and defined as

$$Ad(\mathcal{F}, g_j)(X) = [f, g_j] + \sum_{i=1}^m [g_i, g_j] u_i + \sum_{i=1}^m \sum_{\alpha=0}^n \frac{\partial}{\partial u_i^{(\alpha)}} (g_j) u_i^{(\alpha+1)} \quad (3.10)$$

where  $\mathcal{F}$  is used instead of  $f + \sum_{i=1}^m g_i u_i$ , and  $[\cdot, \cdot]$  denotes the standard Lie bracket. Moreover,  $u_i^{(k)} = \frac{\mathbf{d}^k u_i}{\mathbf{d}t^k}$ . The recursive usage of this bracket is interpreted as the bracketing of vector field  $\mathcal{F}$  with the same vector field  $g_j$ , and is denoted by

$$Ad^k(\mathcal{F}, g_j) = [f, Ad^{k-1}(\mathcal{F}, g_j)] + \sum_{i=1}^m [g_i, Ad^{k-1}(\mathcal{F}, g_j)] u_i + \sum_{i=1}^m \sum_{\alpha=0}^n \frac{\partial}{\partial u_i^{(\alpha)}} (Ad^{k-1}(\mathcal{F}, g_j)) u_i^{(\alpha+1)} \quad (3.11)$$

with  $Ad^0(\mathcal{F}, g_j) = g_j$ .

**Proposition:** Consider the systems governed by Eq. (3.2) with a performance index in the form of Eq. (3.1). The first order necessary conditions of optimality Eq. (3.9) can be rewritten as:

$$\frac{\mathbf{d}^k}{\mathbf{d}t^k} \left( \frac{\partial H}{\partial u_j} \right) = \lambda^T (Ad^k(\mathcal{F}, g_j)) (X); \quad k \in \{0, 1, \dots\}. \quad (3.12)$$

□

It is clear that the first-order necessary conditions for optimality are linear functions of the adjoint states and moreover linear functions of the inputs and their time derivatives.

**Remark:** Notice that

$$\frac{\mathbf{d}}{\mathbf{d}t} (\lambda^T g_j(X)) = \lambda^T \left( [f, g_j] (X) + \sum_{i=1}^m [g_i, g_j] (X) u_i \right)$$

with  $\{i = j\} \implies \{[g_i, g_j](X) = 0\}$ . Then, the first time derivative of the necessary conditions cannot provide the explicit form of  $u_j$ .  $\diamond$

**Proposition:** Consider the system governed by Eq. (3.2) with a performance index in the form of the Eq. (3.1). The optimal state feedback which minimizes the objective function is given by:

$$\det \Lambda_j(X, u) = 0, \quad (3.13)$$

where

$$\Lambda_j(X, u) = \left[ Ad^0(\mathcal{F}, g_j) \dot{=} Ad^1(\mathcal{F}, g_j) \dot{=} \dots \dot{=} Ad^{n-1}(\mathcal{F}, g_j) \right].$$

□

Thus, the Eq. (3.13) results in a state feedback law. Notice that the rank of the matrix generated by the first order necessary conditions allow the characterization of the nature of this state feedback law (dynamic or static). Eq. (3.13) allows three possible cases:

- i)  $\text{rank}(\Lambda_j(X, u)) < n$
- ii)  $\text{rank}(\Lambda_j(X, u)) = n$  and  $u_j$  appears explicitly in Eq. (3.13).
- iii)  $\text{rank}(\Lambda_j(X, u)) = n$  and  $u_j$  do not appear explicitly in Eq. (3.13).

### 3.3.5 The case i) and ii)

Case i) implies that the number of states needed to obtain the optimal input  $u_j$  is equal to the rank  $(\Lambda_j(X, u))$ . And then an adequate choice of the control variables and/or a state can make the analytical solution of this multi-variable optimization problem possible.

Case ii) implies that it is possible to obtain the input  $u_j$  just by solving the Eq. 3.13. Additionally, from the Eq. 3.11 note that if  $k \geq 2$ , in general, the optimal control law will have a dynamic nature. Then, it is necessary to find the initial conditions for this dynamic state feedback law. These initial conditions could be obtained from the off-line solution of the first-order necessary conditions of optimality.

### 3.3.6 The case iii)

Case iii) implies that a singular extremal evolves on the surface  $S_j(X, u) = \det(\Lambda_j(X, u)) = 0$  where the linear independence of

$$Ad^0(\mathcal{F}, g_j) \dot{=} Ad^1(\mathcal{F}, g_j) \dot{=} \dots \dot{=} Ad^{n-1}(\mathcal{F}, g_j)$$

is lost. Then, the optimal solution is equivalent to controlling the system given by the Eq. 3.2 using the input  $u_j$  to the surface  $S_j(X, u) = 0$ . In others words, the optimal problem become a tracking problem.

Feedback linearization is a versatile control technique for nonlinear systems which converts a nonlinear system into a linear system using a nonlinear coordinate transformation and feedback control. Controllers that allow  $y_j(t)$  to track a smooth (at least  $n$ -times differentiable) desired trajectory  $S_j = y_{d_j}(t)$  can be designed by writing the error dynamics,  $e_j(t) = y_j(t) - y_{d_j}(t)$ . It can be seen easily that

$$\frac{d^n e}{dt^n} = \frac{d^n y}{dt^n} - \frac{d^n y_d}{dt^n} = v_j - y_{d_j}^{(n)}(t) \quad (3.14)$$

where  $y_{d_j}^{(n)}(t)$  is the  $n$ -th times derivative of  $y_{d_j}(t)$ . Therefore, a tracking controller can be designed to be:

$$v_j(t) = y_{d_j}^{(n)}(t) - a_{j_0}(z_{j_1} - y_{d_j}) - a_{j_1}(z_{j_2} - \dot{y}_{d_j}) - \dots - a_{j_n}(z_{j_n} - y_{d_j}^{(n-1)})$$

where  $s_{j_n} + a_{j_{n-1}}s_{j_{n-1}} + a_{j_{n-2}}s_{j_{n-2}} + \dots + a_{j_0}$  has all the roots on the left half plane. Consider now the relative degree for each output  $y_j$ . Let  $r_j$  be the smallest relative degree with respect to any input  $u_i$ ,  $i = 1, \dots, m$ . We assume that each  $y_j$  has such a degree. This means that

$$y_j^{(r_j)} = L^{r_j} f h_j(X) + L_g L^{r_j-1} h_j(X) u \quad (3.15)$$

where  $L_g L^{r_j-1} h_j(X) \in \mathbb{R}^{1 \times m}$  with at least one zero element. We can combine these  $m$  equations,  $[y_1, y_2, \dots, y_m]^T = A(X) + B(X)u$  where

$$A(X) = \begin{bmatrix} L^{r_1} f h_1(X) \\ L^{r_2} f h_2(X) \\ \vdots \\ L^{r_m} f h_m(X) \end{bmatrix}; \quad B(X) = \begin{bmatrix} L_g L^{r_1-1} h_1(X) \\ L_g L^{r_2-1} h_2(X) \\ \vdots \\ L_g L^{r_m-1} h_m(X) \end{bmatrix}$$

Notice that  $A(X) \in \mathbb{R}^m$  and  $B(X) \in \mathbb{R}^{m \times m}$ . The feedback linearization can be easily extended to MIMO systems if  $B(X)$  is invertible. In that case, we can design

$$u = B(X)^{-1}[-A(X) + v] \quad (3.16)$$

where  $v(t) \in \mathbb{R}^m$  is an auxiliary input. This generate a set of decoupled and linear input-output systems, for  $k = 1, \dots, m$ ,  $y_j^{(r_j)} = v_j$ . A MIMO system is said to have a vector relative degree  $(r_1 \ r_2 \ \dots \ r_m)$  if the individual outputs  $y_j$  have a relative degree  $r_j$  and  $B(x)$  is invertible. It turns out that if  $B(x)$  is invertible, then the output functions and the derivatives of the output functions can be used as independent coordinate functions. If

$$\sum_{j=1}^m r_j = n \quad (3.17)$$

then, there will not be any internal dynamics. Note, that it is important to choose an appropriate output  $y_j = h_j(X)$ .

The function  $\lambda^T g_i(X)$  is called the switching function. This function vanishes over the singular time interval. Outside the singular interval, the manipulated input,  $u_i$ , takes either its maximum value,  $u_{i_{max}}$ , or its minimum value,  $u_{i_{min}}$ , depending on the sign of the switching function. When  $\lambda^T g_i(X) < 0$ ,  $u_i = u_{i_{max}}$  and  $\lambda^T g_i(X) > 0$ ,  $u_i = u_{i_{min}}$ . The times at which the input switches from singular to nonsingular interval and *vice versa* is called the switching time. The necessary condition for optimality can be verified by integrating the adjoint equation ( $\dot{\lambda}^T g_i(X) = 0$ ) backwards in time. Note that the final condition for the backward integration is given by the Eq. (3.5), which involves the performance index.

### 3.4 The Ideal Model

Biomass production is necessary in many biochemical plants, for example the production of baker's yeast and as a prerequisite for production of food additives and recombinant proteins. For biomass production three basic modes of operation are possible: continuous, batch or fed-batch. Continuous operation is usually not the most productive option as sufficient exhaustion of the enriched feed can only be realized at rather low feed rates. In batch and fed-batch modes, a part of the culture biomass can be left in the reactor for the next batch, enabling a cyclic operation of repeated batches or fed-batches.

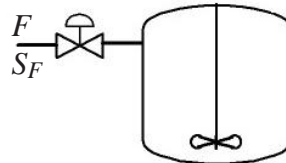


Figure 3.1: Schematic of the culture vessel for cell production

This chapter considers the problem of following an OBBOM for a bioreactor in the presence of parametric uncertainties associated with the feed substrate concentration and the cell death rate. The optimal operations model is the optimal trajectory that the feed flow and the substrate concentration should have for the nominal cultivation to be optimal.

In order to study the scheme for bioreactors, a relatively simple mathematical model is selected. The system consists of a mixing tank, where two streams are fed, one with nutrients at high concentration and the other with a solvent including supplementary components. The reactor is

assumed perfectly stirred. An unstructured biomass growth rate equation with substrate inhibition kinetics is chosen. Only a limited set of states is available for measuring.

In the fed-batch fermenter for cell production the substrate is converted by biomass into additional biomass. It is assumed that the reactor is ideally mixed. The biomass and substrate are represented by their concentration in the culture, denoted by  $X_1$  and  $X_2$ , respectively. The general unstructured mass balances for the well-mixed bioreactor can be represented by the following equations for cells, substrate and reactor volume ( $X_3$ ):

$$\dot{X}_1 = X_1 \mu(X_2) - \frac{X_1 F}{X_3} \quad (3.18)$$

$$\dot{X}_2 = -\frac{X_1}{Y} \mu(X_2) + \frac{F}{X_3} (S_F - X_2) \quad (3.19)$$

$$\dot{X}_3 = F \quad (3.20)$$

where  $Y$  is the biomass yield and  $\mu(X_2)$  is the specific growth rate. The growth rate relates the change in biomass concentration to the substrate concentration. Two types of relationships for  $\mu(X_2)$  are commonly used: the substrate saturation model (Monod Equation) and substrate inhibition model (Haldane Equation). The substrate inhibited growth can be described by

$$\mu(X_2) = \frac{\mu_{max} X_2}{K_1 + X_2 + K_2 X_2^2} \quad (3.21)$$

where  $K_1$  is the saturation or Monod constant,  $K_2$  is the inhibition constant and  $\mu_{max}$  is the maximum specific growth rate. The value of  $K_1$  expresses the affinity of biomass for substrate. The Monod growth kinetics can be considered as a special case of the substrate inhibition kinetics with  $K_2 = 0$  when the inhibition term is vanished.

The problem is to maximize the production of cell mass for the above dynamic system by finding the optimal filling policy  $F(t)$  and the optimal substrate concentration in the inlet stream  $S_F(t)$ . In order to apply the results developed in this chapter the optimization problem is formulated as follows. Minimize the performance index

$$J = -X_1|_{t=t_f} \quad (3.22)$$

subject to the dynamic  $\dot{X} = f(X) + g_1(X, u_2)u_1$  where  $\dot{X} = [X_1 \ X_2 \ X_3]$ ,

$$f(X) = \begin{bmatrix} X_1 \mu(X_2) \\ -\frac{X_1}{Y} \mu(X_2) \\ 0 \end{bmatrix}, \quad g_1(X, u[2]) = \begin{bmatrix} -\frac{X_1}{X_3} \\ \frac{u_2 - X_2}{X_3} \\ 1 \end{bmatrix} \quad (3.23)$$

$u_1 = F$  and  $u_2 = S_F$ . And we can set  $g_2(X, u_1) = \frac{\partial(g_1(X, u_2)u_1)}{\partial u_2}$ . It is important to remark the possibility to choose a different performance index. As it is mentioned before the performance index will be used to find the switching time, therefore the analysis is exactly the same and only the condition for the backward integration of the adjoint equation,  $\lambda^T g_i(X)$ . Also notice, that for this particular case we have  $n = 3$  and therefore

$$\mathbf{rank} \left[ Ad^0(\mathcal{F}, g_1) : Ad^1(\mathcal{F}, g_1) : Ad^2(\mathcal{F}, g_1) \right] = 2 \quad (3.24)$$

$$\mathbf{rank} \left[ Ad^0(\mathcal{F}, g_2) : Ad^1(\mathcal{F}, g_2) : Ad^2(\mathcal{F}, g_2) \right] = 2 \quad (3.25)$$

The previous conditions could be interpreted as the existence of redundant equation in the system given by the Eqs. 3.18-3.20. from the point of view of the optimization problem.

An alternative to avoid this inconvenience was pointed out by [29] for the single input case. That is, taking the dilution rate as one of the control variable reduces the dimension of the system by making use of the total mass balance which therefore is unnecessary for the solution of the optimization problem. Thus, the model given by the Eqs. 3.18-3.20 can be rewritten as

$$\dot{X}_1 = X_1 \mu(X_2) - X_1 D \quad (3.26)$$

$$\dot{X}_2 = -\frac{X_1}{Y} \mu(X_2) + D(S_F - X_2) \quad (3.27)$$

$$\dot{X}_3 = X_3 D. \quad (3.28)$$

Hence, it is possible to consider just the two first equations by taking as the control inputs  $u_1 = D$  and  $u_2 = S_F$ . This leads to the following reduced model  $\dot{X} = f(X) + g_1(X, u_2)u_1$ , where now  $\dot{X} = [X_1 \ X_2]$  and

$$f(X) = \begin{bmatrix} X_1 \mu(X_2) \\ -\frac{X_1}{Y} \mu(X_2) \end{bmatrix}, \quad g_1(X, u_2) = \begin{bmatrix} -X_1 \\ u_2 - X_2 \end{bmatrix}. \quad (3.29)$$

We can again define  $g_2(X, u_1) = \frac{\partial(g_1(X, u_2)u_1)}{\partial u_2}$ . Thus, the necessary conditions for optimality are given by

$$\lambda^T \left[ g_1 : [f, g_1] \right] = \lambda^T \Lambda_1(X, u) \quad (3.30)$$

$$\lambda^T \left[ g_2 : [f, g_2] \right] = \lambda^T \Lambda_2(X, u). \quad (3.31)$$

The Eqs. (3.30)-(3.31) are related to the necessary condition for the input  $u_1$  and  $u_2$ , respectively. Neither of the two equations provides an explicit function for the corresponding inputs.

$$\Lambda_1(X, u) = \begin{bmatrix} -X_1 & -X_1 \frac{\partial \mu(X_2)}{\partial X_2} (u_2 - X_2) \\ u_2 - X_2 & \frac{X_1}{Y} \frac{\partial \mu(X_2)}{\partial X_2} (u_2 - X_2) \end{bmatrix} \quad (3.32)$$

$$\Lambda_2(X, u) = \begin{bmatrix} 0 & -X_1 \frac{\partial \mu(X_2)}{\partial X_2} u_1 \\ u_1 & \frac{X_1}{Y} \frac{\partial \mu(X_2)}{\partial X_2} u_1 \end{bmatrix}. \quad (3.33)$$

As discussed in the previous section under case *iii*), in this case, the optimal state trajectories must lie on the surfaces:

$$S_1 = \det(\Lambda_1(X, u)) = \frac{X_1}{Y} \frac{\partial \mu(X_2)}{\partial X_2} (u_2 - X_2) (-X_1 + Y(u_2 - X_2)) = 0 \quad (3.34)$$

$$S_2 = \det(\Lambda_2(X, u)) = X_1 \frac{\partial \mu(X_2)}{\partial X_2} u_1^2 = 0. \quad (3.35)$$

In general, the optimization problem becomes a tracking problem. Since  $X_1(t) \neq 0$  and  $X_2(t) \neq u_2(t)$ , from the Eq. (3.34) we can note that  $S_1 = 0$  implies

$$\frac{\partial \mu(x_2)}{\partial x_2} = -\frac{\mu_{\max} (K_2 X_2^2 - K_1)}{(K_2 X_2^2 + X_2 + K_1)^2} = 0. \quad (3.36)$$

The value of  $X_2$  that fulfils the previous equation is clearly:

$$X_2^* = \sqrt{\frac{K_1}{K_2}}. \quad (3.37)$$

Now, the problem is to find an admissible control law  $u_1$  such that  $X_2 = X_2^*$ . It is very important to note that the control law  $u_1$  that keeps  $S_1 = 0$  also will ensure the second necessary condition  $S_2 = 0$  for all admissible value of  $u_2$ .

In [41] it is shown that, if the dilution rate is constant, this simplified model has three equilibrium points. The expression for the substrate values at the equilibrium points are given by:

$$X_{2Eq1} = S_F; \quad X_{2Eq2} = -\omega; \quad X_{2Eq3} = -\phi; \quad (3.38)$$

where

$$\phi = \frac{D - \mu_{\max}}{2DK_2} + \sqrt{\left(\frac{D - \mu_{\max}}{2DK_2}\right)^2 - \frac{K_1}{K_2}} \quad (3.39)$$

$$\omega = \frac{D - \mu_{\max}}{2DK_2} - \sqrt{\left(\frac{D - \mu_{\max}}{2DK_2}\right)^2 - \frac{K_1}{K_2}}. \quad (3.40)$$



As the dilution rate increases the two equilibrium points  $X_{Eq2}$  and  $X_{Eq3}$  collide. This kind of bifurcation is called a fold bifurcation. The value of the bifurcation parameter at which the two equilibria collide is called the dilution rate bifurcation point and is given by the following expression:

$$D^* = \frac{\mu_{\max}}{1 + 2\sqrt{K_1 K_2}} \quad (3.41)$$

In other words, the above dilution rate value make  $\omega = \phi$  and therefore

$$X_{2Eq2} = X_{2Eq3} = X_2^* = \sqrt{\frac{K_1}{K_2}} \quad (3.42)$$

This point is locally stable. Then, when we achieve the value of  $X_2 = X_2^*$ , the biomass concentration is given by:

$$X_1^* = -\frac{Y(-u_2(2K_1K_2 + \sqrt{K_1K_2}) + 2\sqrt{K_1K_2}K_1 + K_1)}{2K_1K_2 + \sqrt{K_1K_2}}$$

$$X_1^* = u_2 \frac{Y(2K_1K_2 + \sqrt{K_1K_2})}{2K_1K_2 + \sqrt{K_1K_2}} - \frac{Y(2\sqrt{K_1K_2}K_1 + K_1)}{2K_1K_2 + \sqrt{K_1K_2}}$$

$$X_1^* = u_2 Y - \frac{K_1 Y}{\sqrt{K_1 K_2}} = u_2 Y - Y \sqrt{\frac{K_1}{K_2}}. \quad (3.43)$$

From the latter expression we can note that

$$u_2^\oplus = \sqrt{\frac{K_1}{K_2}} = X_2^*$$

is the value of the substrate concentration that produces zero biomass concentration. That is clear in the sense that the inlet substrate concentration would be equal to the outlet substrate concentration. On the other hand, once given this constraint on the biomass, the Eq. (3.43) provides the optimal substrate concentration.

$$u_2^* = \frac{X I_{\max}}{Y} + \sqrt{\frac{K_1}{K_2}}. \quad (3.44)$$

The previous results show that it is possible to achieve the control goal for all  $u_2 > X_2^*$ . Then a proper initial condition and open-loop operation can provide the necessary condition for optimality. To avoid be limited to a set of initial conditions, we can design a closed-loop control law.

### The Tracking Problem

In the previous subsection it was shown that it is only sufficient to design a control law for  $u_1$ , and  $u_2$  is given by the constraint on biomass and Eq. (3.43). Consider the nonlinear time invariant system:

$$\begin{aligned}\dot{X} &= f(X) + g(X)u \\ y &= h(X)\end{aligned}$$

where the output that allows the exact feedback linearization is given by

$$h(X) = \frac{X_2 - S_F}{X_1}$$

Notice that under this output, we have the relationships

$$\begin{aligned}L_g h(X) &= 0 \\ L_g L_f h(X) &= h(X) \frac{\partial \mu(X_2)}{\partial X_2} \left( \frac{1}{Y} + h(X) \right) X_1 \\ L_f^2 h(X) &= \left( \frac{\mu^2(X)}{X_1} + \frac{\mu(X)}{Y} \frac{\partial \mu(X_2)}{\partial X_2} \right) \left( \frac{1}{Y} + h(X) \right) X_1\end{aligned}$$

implying that the relative degree of the output  $h(X)$  is  $r = 2$ . Therefore, the controller is given by

$$u_1 = \alpha(-\beta + v_1), \quad (3.45)$$

where

$$\alpha = \frac{1}{L_g L_f h(X)} \quad (3.46)$$

$$\beta = L_f^2 h(X) \quad (3.47)$$

and  $v_1$  is the dynamics that is desirable to be imposed.

$$v_1 = -a_{10}(h(X) - \Lambda) - a_{20}(z_2);$$

and

$$\Lambda = \frac{X_2^* - u_2^*}{X_1^*}, \quad z_2 = -\mu(X_2) \left( h(X) + \frac{1}{Y} \right).$$

The feedback law given by the Eq. (3.45) guarantees the first-order necessary condition for optimality. Since, in general, a bioreactor starts with a high substrate concentration, there will be an initial nonsingular phase where no substrate is fed. The switch to the singular phase occurs when  $X_2$  reaches the value  $X_2^*$ . Therefore, the closed-loop system is that one which has the trajectory of the feed flow and the substrate concentration should track in order to get the optimal cultivation. Thus, hereafter this system will be considered as the master system.

In the next section a more *Realistic Model* is considered, by means of which a perturbed fed-batch bioreactor model will be studied and taken as the slave system.

### 3.5 The Realistic Model

The configuration studied here for the fed-batch bioreactor of cell production is shown schematically in Figure 3.2.

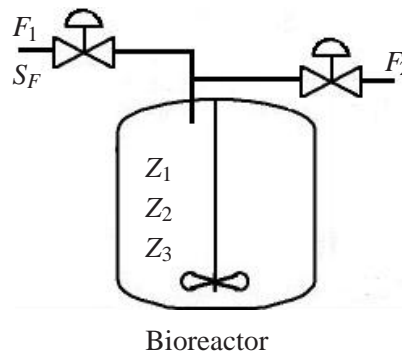


Figure 3.2: The figure shows the system, which consists of a tank reactor, where two streams are fed.  $F_1$  with substrate at high concentration and  $F_2$  with a solvent including supplementary components. The reactor is assumed perfectly stirred.

The assumptions for the analysis are: (i) open-loop operation, (ii) all model parameters and physicochemical properties are real and finite, but unknown, (iii) the fermentation occurs in a perfectly mixed continuous bioreactor operated isothermally. Under such assumptions, the model is given by the following differential equations:

$$\frac{dZ_1}{dt} = Z_1 (\mu(Z_2) - r_d) - \frac{Z_1}{Z_3} F_1 - \frac{Z_1}{Z_3} F_2 \quad (3.48)$$

$$\frac{dZ_2}{dt} = -\frac{Z_1}{Y} \mu(Z_2) + \frac{S_f - Z_2}{Z_3} F_1 - \frac{Z_2}{Z_3} F_2 \quad (3.49)$$

$$\frac{dZ_3}{dt} = F_1 + F_2 \quad (3.50)$$

where  $Z_1$  represent the Biomass conc. [g/l],  $Z_2$  is the Substrate conc. [g/l],  $Z_3$  is the volume [l] in the Bioreactor.  $F_1$  is the feed flow rate [l/h] with the limiting substrate, and  $F_2$  is the second feed flow rate [l/h] free of the limiting substrate.

Table 3.1: Operating variables and parameters of the fermentation process model.

Symbol	Meaning	Value	Units
$S_f$	Substrate feed conc.	20	[g/l]
$Y$	Yield coefficient	0.5	—
$\mu_{max}$	Maximal growth rate	1	[l/h]
$K_1$	Saturation parameter	0.03	[g/l]
$K_2$	Inhibition parameter	0.5	[l/g]
$F_1$	Feed Flow rate (input 1)	—	[l/h]
$F_2$	Feed Flow rate (input 2)	—	[l/h]

The inlet substrate concentration and the cell death rate will be considered unknown. Therefore, a robust control structure should be designed to make that this model to follow the master system in spite of exogenous disturbances.

For the sake of simplicity, the model given by the Eqs. (3.48)-(3.50) can be rewritten as

$$\frac{dZ}{dt} = \hat{f}(Z) + \sum_{i=1}^2 \hat{g}_i(Z) \hat{u}_i$$

where

$$\hat{f}(Z) = \begin{bmatrix} Z_1 (\mu(Z_2) - r_d) \\ -\frac{Z_1}{Y} \mu(Z_2) \\ 0 \end{bmatrix}, \quad g_1(Z) = \begin{bmatrix} -\frac{Z_1}{Z_3} \\ \frac{S_f - Z_2}{Z_3} \\ 1 \end{bmatrix}, \quad g_2(Z) = \begin{bmatrix} -\frac{Z_1}{Z_3} \\ -\frac{Z_2}{Z_3} \\ 1 \end{bmatrix}.$$

In the next section, the strategy to follow the Optimal Batch Operation Model will be designed. The main idea is based on a master slave synchronization using a robust control law based on differential geometry. Additionally, a reachability analysis will be performed and the exact feedback linearization will be studied in order to show the advantages of using two inputs.

## 3.6 Following the Optimal Batch Operation Model

### 3.6.1 Reachability Analysis

Reachability analysis is an important tool in analysis and synthesis of control systems. It refers to the problem of knowing if it is possible to steer a dynamical system from the point  $z^0$  to  $z^1$ . If this steering is possible, we say that  $z^1$  is reachable from  $z^0$ . Reachability analysis of the SISO fed-batch bioreactor has received some attention previously ([39]). In this section, we consider the reachability analysis of the MIMO case where both feed concentration and flow rate are inputs. The aim is to provide an understanding of the advantages of employing two inputs.

The reachability can be investigated (at least locally) by means of the differential geometric methods. The methodology to determine, locally, the reachability properties from a given affine systems was clearly established in [38]. The following algorithm constructs the reachability distribution given  $m + 1$  vector fields  $\{\tau_1, \tau_2, \dots, \tau_q\} = \{f, \hat{g}_1, \dots, \hat{g}_m\}$ .

$$\begin{aligned}\Delta_0 &= \text{span}\{\hat{g}_1 \dots \hat{g}_m\} \\ \Delta_i &= \Delta_{k-1} + \sum_{i=1}^q [\tau_i, \Delta_{k-1}]\end{aligned}$$

for  $i = 0, 1, \dots, n-1$ , where  $n$  is the number of states. The application of the above methodology to the systems given by the Eqs. (3.48)-(3.50), for which  $n = 3$ , constructs the following sequence

$$\begin{aligned}\Delta_0 &= \text{span}\{\hat{g}_1, \hat{g}_2\} \\ \Delta_1 &= \text{span}\{\hat{g}_1, \hat{g}_2, [\hat{f}, \hat{g}_1], [\hat{f}, \hat{g}_2], [\hat{g}_1, \hat{g}_2], [\hat{g}_2, \hat{g}_1]\}\end{aligned}$$

and

$$\begin{aligned}\Delta_2 = \Delta_1 + \text{span}\{ & [\hat{f}, \hat{g}_1], [\hat{f}, \hat{g}_2], [\hat{f}, [\hat{f}, \hat{g}_1]], [\hat{f}, [\hat{f}, \hat{g}_2]], [\hat{f}, [\hat{g}_1, \hat{g}_2]], [\hat{f}, [\hat{g}_2, \hat{g}_1]] \\ & , [\hat{g}_1, \hat{g}_2], [\hat{g}_1, [\hat{f}, \hat{g}_1]], [\hat{g}_1, [\hat{f}, \hat{g}_2]], [\hat{g}_1, [\hat{g}_1, \hat{g}_2]], [\hat{g}_1, [\hat{g}_2, \hat{g}_1]] \\ & , [\hat{g}_2, \hat{g}_1], [\hat{g}_2, [\hat{f}, \hat{g}_1]], [\hat{g}_2, [\hat{f}, \hat{g}_2]], [\hat{g}_2, [\hat{g}_1, \hat{g}_2]], [\hat{g}_2, [\hat{g}_2, \hat{g}_1]]\}\end{aligned}$$

$$\hat{f}(Z) = \begin{bmatrix} Z_1(\mu(Z_2) - r_d) \\ -\frac{Z_1}{Y}\mu(Z_2) \\ 0 \end{bmatrix}, \quad \hat{g}_1(Z) = \begin{bmatrix} \hat{g}_{21} \\ \frac{S_f}{Z_3} - \hat{g}_{22} \\ 1 \end{bmatrix}, \quad \hat{g}_2(Z) = \begin{bmatrix} \hat{g}_{21} \\ \hat{g}_{22} \\ 1 \end{bmatrix} \quad (3.51)$$

where  $\hat{g}_{21}$  and  $\hat{g}_{22}$  represent the first and second input of the vector  $\hat{g}_2$  respectively, and  $\hat{g}_2$  was previously defined in Eq. (3.51). From the above representation of  $\hat{g}_1$  and  $\hat{g}_2$  it is clear that

$$\text{rank}(\Delta_0) = \text{rank}[\hat{g}_1, \hat{g}_2] = 2 \quad (3.52)$$

and it is possible prove that

$$\frac{\partial \hat{g}_1(Z)}{\partial Z} = \begin{bmatrix} -Z_3^{-1} & 0 & \frac{Z_1}{Z_3^2} \\ 0 & -Z_3^{-1} & -\frac{S_f - Z_2}{Z_3^2} \\ 0 & 0 & 0 \end{bmatrix}; \quad \frac{\partial \hat{g}_2(Z)}{\partial Z} = \begin{bmatrix} -Z_3^{-1} & 0 & \frac{Z_1}{Z_3^2} \\ 0 & -Z_3^{-1} & \frac{Z_2}{Z_3^2} \\ 0 & 0 & 0 \end{bmatrix}.$$

Therefore

$$\frac{\partial \hat{g}_2(Z)}{\partial Z} \hat{g}_1(Z) = \frac{\partial \hat{g}_1(Z)}{\partial Z} \hat{g}_2(Z) = \left[ 2 \frac{Z_1}{Z_3^2}, -\frac{S_f - Z_2}{Z_3^2} + \frac{Z_2}{Z_3^2}, 0 \right]^T. \quad (3.53)$$

On the other hand, we have

$$\frac{\partial \hat{f}(Z)}{\partial Z} = \begin{bmatrix} \mu(Z_2) - r_d & \frac{\partial \mu(Z_2)}{\partial Z_2} Z_1 & 0 \\ \frac{\mu(Z_2)}{Y_s} & -\frac{\partial \mu(Z_2)}{\partial Z_2} \left( \frac{Z_1}{Y_s} \right) & 0 \\ 0 & 0 & 0 \end{bmatrix}. \quad (3.54)$$

Then

$$\begin{aligned} [f, \hat{g}_1] &= \left[ \frac{\partial \mu(Z_2)}{\partial Z_2} \left( \frac{Z_1(Z_2 - S_f)}{Z_3} \right), -\frac{\partial \mu(Z_2)}{\partial Z_2} \left( \frac{Z_1 Z_2}{Y_s Z_3} \right), 0 \right] \\ [f, \hat{g}_2] &= \left[ \frac{\partial \mu(Z_2)}{\partial Z_2} \left( \frac{Z_1 Z_2}{Z_3} \right), -\frac{\partial \mu(Z_2)}{\partial Z_2} \left( \frac{Z_1 Z_2}{Y_s X_3} \right), 0 \right] \end{aligned}$$

yielding

$$\text{rank}(\Delta_1) = \text{rank} [\hat{g}_1, \hat{g}_2, [\hat{f}, \hat{g}_1], [\hat{f}, \hat{g}_2]] = 3.$$

Since the systems has 3 states, to prove that  $\text{rank}(\Delta_2) = 3$  is trivial.

In [39] it is shown that the difficulties for controlling fed-batch fermentation processes are primarily caused by the fact that the rank of the reachability distribution is always less than the number of state variables. In practical terms this means that if we want to reach a predefined state at the end of a batch and only the inlet flow rate of the substrate solution of a given concentration is manipulated then it is necessary to choose an appropriate initial state in order to achieve the goal of matching the desired set of states.

Here it is shown that the usage of two inputs overcomes the aforementioned problems. The inclusion of the extra input variable guarantees the rank of the reachability distribution. The rigorous nonlinear analysis presented above provides a formal explanation for the advantages of employing an additional input.

### 3.6.2 Studying the Exact Feedback Linearization Problem

The exact linearization ([45]) with the corresponding geometric approach ([38]) is one of the most effective and fundamental techniques in the field of nonlinear control. This technique is widely used for actual nonlinear plants and plays an important role in nonlinear control systems theory.

Exact linearization is a method to transform a nonlinear system into a linear one, and it is usually used for stabilization and tracking control of nonlinear systems with a linear compensator for the linearized system. In this design procedure, a solution of a set of partial differential equations is used to obtain the linearizing transformation.

The necessary and sufficient condition to achieve the exact linearization are given in the Theorem 3.6.2. To present the theorem the following distributions are defined:

$$\begin{aligned} G_0 &= \text{span} \{ \hat{g}_1 \dots \hat{g}_m \} \\ G_1 &= \text{span} \{ \hat{g}_1 \dots \hat{g}_m, \text{ad}_f \hat{g}_1 \dots \text{ad}_f \hat{g}_m \} \\ &\dots \\ G_i &= \text{span} \left\{ \text{ad}_f^k \hat{g}_j : 0 \leq k \leq i, 1 \leq j \leq m \right\} \end{aligned}$$

for  $i = 0, 1, \dots, n-1$ .

**Theorem** suppose that the matrix  $g(z^0)$  has rank  $m$ . Then, the state space exact linearization is solvable if and only if

- (i) for each  $0 \leq i \leq n-1$ , the distribution  $G_i$  has constant dimension near  $z^0$ ;
- (ii) the distribution  $G_{n-1}$  has dimension  $n$ ;
- (iii) for each  $0 \leq i \leq n-2$ , the distribution  $G_i$  is involutive.

i.e., there exist  $m$  real-valued function  $\lambda_1(x), \lambda_1(z), \dots, \lambda_m(z)$  defined on a neighborhood  $U$  of  $z^0$ , such that the system of the form

$$\begin{aligned} \dot{x} &= f(x) + \sum_{i=0}^m g_i(x) u_i \\ y_i &= \lambda_i(x) \end{aligned}$$

has a vector relative degree  $\{r_1, \dots, r_m\}$  at  $x^0$ , with  $r_1 + r_2 + \dots + r_m = n$ . ■

In others words, the main issue, in order to achieve the exact feedback linearization, is to find the solutions  $\lambda_1(Z), \lambda_1(Z), \dots, \lambda_m(Z)$  of equations of the form

$$L_{g_j} L^k f \lambda_i(Z) = 0 \quad \text{for all } 0 \leq k \leq r_i - 2, 1 \leq j \leq m. \quad (3.55)$$

In the case treated here the previous equations have the following explicit form:

$$-\frac{\partial \lambda_1(Z)}{\partial Z_1} \left( \frac{Z_1}{Z_3} \right) + \frac{\partial \lambda_1(Z)}{\partial Z_2} \left( \frac{S_f - Z_2}{Z_3} \right) + \frac{\partial \lambda_1(Z)}{\partial Z_3} = 0 \quad (3.56)$$

$$-\frac{\partial \lambda_1(Z)}{\partial Z_1} \left( \frac{Z_1}{Z_3} \right) - \frac{\partial \lambda_1(Z)}{\partial Z_2} \left( \frac{Z_2}{Z_3} \right) + \frac{\partial \lambda_1(Z)}{\partial Z_3} = 0 \quad (3.57)$$

In this case, their solutions are given by

$$\lambda_1(Z) = \ell \left( \frac{Z_3}{Z_1} \right) \quad (3.58)$$

$$\lambda_2(Z) = \ell(Z_2) \quad (3.59)$$

where  $\ell : \mathbb{R} \rightarrow \mathbb{R}$  could be any function.

Therefore,  $r_1 + r_2 = 3$  and the relative degree matrix is given by

$$Rdm(X) = \begin{bmatrix} X_3 X_1 \frac{\partial \mu(X_2)}{\partial X_2} (S_f - X_2) & X_3 X_1 \frac{\partial \mu(X_2)}{\partial X_2} X_2 \\ X_3 & X_3 \end{bmatrix}. \quad (3.60)$$

Following Eq. (3.15), the system in the new coordinates is given by

$$\begin{bmatrix} \dot{y}_1 \\ \dot{y}_2 \end{bmatrix} = \begin{bmatrix} L_f^2 h_1(X) \\ L_f h_2(X) \end{bmatrix} + \begin{bmatrix} L_{g_1} L_f h_1(X) & L_{g_2} L_f h_1(X) h_1(X) \\ L_{g_1} h_2(X) & L_{g_2} h_2(X) \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}$$

and we can appreciate that

$$B^{-1}(X) = \begin{bmatrix} \frac{1}{X_1 X_3 S_f} \left( \frac{\partial \mu(X_2)}{\partial X_2} \right)^{-1} & \frac{X_2}{S_f X_3} \\ -\frac{1}{X_1 X_3 S_f} \left( \frac{\partial \mu(X_2)}{\partial X_2} \right)^{-1} & \frac{S_f - X_2}{S_f X_3} \end{bmatrix}. \quad (3.61)$$

Thus, the conditions imposed by Theorem 3.6.2 are fulfilled.



In the two previous works of [40], the expression for the output that achieves the exact linearization is a function of the states and the inlet substrate concentration ( $S_F$ ). But if we take into account the fact that to know the exact value of  $S_F$  with high precision is very difficult and moreover this value could be also subject to continuous perturbations. Therefore, we can note that the application of the exact linearization in the SISO case is practically impossible.

In general terms, we can say that this is an additional advantage in the usage of two input for control purposes. This is due to the fact that the outputs of the system that achieve the exact linearization only depend on the states and not on the parameters, see Eqs. (3.58)-(3.59). Although, for the particular control objectives here treated we can note that the inversion of the matrix  $B^{-1}(X)$  depends on the inverse of  $\frac{\partial \mu(X_2)}{\partial X_2}$ , equations that we know will be equal to zero at the final time. Therefore, the exact linearization is not an alternative to follow the optimal batch operation model. However, this problem could be also relevant for others control purposes.

### 3.6.3 Robust Nonlinear Synchronization

#### Synchronization

Synchronization is often categorized on the basis of whether the coupling mechanism is unidirectional or bidirectional. Stable synchronization with unidirectional coupling has been called master-slave synchronization (although [44] shows that unidirectionally and bidirectionally coupled synchronized systems are locally equivalent).

Here, a master-slave configuration is used to follow an optimal batch operations model for a Bioreactor. That is, the optimal operations model is considered as the master system which defines the trajectory that the feed flow rate and the substrate concentration should follow in order to achieve an optimal cultivation, whereas the "real" bioreactor (the one with unknown dynamics and perturbations) is considered as the slave. The advantages of this idea are that we can easily perform the optimization for a simplified model, and then just design the controller such that the real bioreactor is synchronized with the optimized one in spite of unknown dynamics and perturbations. Hereafter both the master and the slave will be considered as deterministic finite dimensional dynamical system.

In order to allow a rigorous treatment of this idea, it is necessary to introduce a very clear definition of synchronization. Colloquially, synchronization is the concurrence of events or motions with respect to time. From this intuitive definition Ref. [43] points out that synchronization requires the following tasks:

- Separating the dynamics of a large dynamical system into the dynamics of subsystems.
- Measuring properties of the subsystems.

- Comparing properties of the subsystems.
- Determining whether the properties agree in time.

If the properties agree then the systems are synchronized. To introduce a rigorous definition of Synchronization, assume that a large stationary deterministic finite-dimensional dynamical system is divided into two subsystems

$$\frac{dX}{dt} = \mathbf{F}_1(X, Z); \quad \frac{dZ}{dt} = \mathbf{F}_2(X, Z), \quad (3.62)$$

where  $X \in \mathbb{R}^{d_1}$  and  $Y \in \mathbb{R}^{d_2}$  are vectors which may have different dimensions. The phase space and vector field of the large system is formed from the product of the two smaller phase spaces and vector fields.

Let us consider the trajectory of the large dynamical system, given by Eq. (3.62), with the initial condition  $Y_0 = [X_0, Z_0] \in \mathbb{R}^{d_1} \otimes \mathbb{R}^{d_2}$ . And the curves  $\phi_X(Z_0)$  and  $\phi_Y(Z_0)$  are the projections of the components  $X$  and  $Z$  respectively. We say that  $\phi_X(Y_0)$  and  $\phi_Z(Y_0)$  are trajectories of the first and second subsystem of Eq. (3.62).

Let  $\chi$  denote the space of all trajectories of the first subsystem, and consider the function  $G_X : \chi \otimes \mathbb{R} \rightarrow \mathbb{R}^k$  which is not identically zero. The first  $\mathbb{R}$  represent the time, and is included such that  $G_X$  may make explicit reference to time. The function  $G_X$  is considered a *property* of the first subsystem. The image of  $[\phi_X(Y_0), t] \in \chi \otimes \mathbb{R}$  under  $G_X$  is the result of *measuring the property* of the first subsystem, and will be denoted by  $G(X) \in \mathbb{R}^k$ . Similar definitions can be made for the second subsystem.

Here, it is assumed that this property phase belongs to the space of the subsystem, therefore measuring the property means determining the values for the coordinates. Therefore the property depends on time; In particular the property being measured, and  $G(X) = X(t)$  provides the values of the measurement. Two measurements agree in time if and only if the time independent function  $H : \mathbb{R}^k \otimes \mathbb{R}^k \rightarrow \mathbb{R}^k$  that compares the measurement properties is such that  $H[G(X), G(Y)] = 0$ .

**Definition** ([43]) The subsystems in Eq. 3.62 are synchronized with respect to the properties  $G_X$  and  $G_Y$  if there is a time independent mapping  $H : \mathbb{R}^k \otimes \mathbb{R}^k \rightarrow \mathbb{R}^k$  such that

$$\|H[G(X), G(Z)]\| = 0 \quad (3.63)$$

holds on all trajectories.  $\square$

Here  $\|\cdot\|$  is any norm. Based upon the previous definition the following practical definition is introduced.

**Definition**[Practical Synchronization] The subsystems in Eq. 3.62 are synchronized with respect to the properties  $G_X$  and  $G_Y$  if there is a time independent mapping  $H : R^k \otimes R^k \rightarrow R^k$  such that

$$\|H[G(X), G(Z)]\| \leq \varepsilon \quad (3.64)$$

holds on all trajectories.  $\square$

Here,  $\varepsilon$  is the maximum tolerable separation between the subsystem trajectories. In particular, the following comparison function here used:

$$H[G(X), G(Z)] = G(X) - G(Z). \quad (3.65)$$

### 3.6.4 Robust Control Law

Here we consider a robust control problem for a class of MIMO nonlinear systems which are minimum phase and of relative degree  $\{1, 1, \dots, 1\}$ . First we consider the problem of designing a *high gain feedback nonlinear controller* that will render the system asymptotically stable under any unknown dynamics and perturbations. Next we pay attention to the *sliding mode control approach* for the same problem. With the differential geometric method of nonlinear systems, [42] reviewed some important results about the passivity of nonlinear systems and obtained that, under mild regularity assumptions, an affine nonlinear system is locally feedback equivalent to a passive system if, and only if, the system has relative degree:  $\{1, \dots, 1\}$  and is weakly minimum phase [[42], Theorem 4.7].

Let us consider the following uncertain MIMO process system as

$$\dot{z} = f(z) + \Delta f(z) + \left( \sum_{i=1}^m g_i(z) + \sum_{i=1}^m \Delta g_i(z) \right) u \quad (3.66)$$

$$y = h(z) \quad (3.67)$$

where  $z \in \mathbb{R}^n$ ,  $u \in \mathbb{R}^m$ ,  $y \in \mathbb{R}^m$ , and  $f(\cdot)$ ,  $\Delta f(\cdot)$ ,  $g_i(\cdot)$ ,  $\Delta g_i(\cdot)$  are smooth vector fields on an open set  $U \in \mathbb{R}^n$ .

It is assumed that the state vector is measured or estimated from available measurements. Moreover, let us take the unperturbed form of Eqs. (3.66-3.67) as the **nominal system** i.e.:

$$\dot{z} = f(z) + \sum_{i=1}^m g_i(z) u_i \quad (3.68)$$

$$y = h(z) \quad (3.69)$$

Consider the so called **strong relative degree** for each output  $y_i$ . Let  $\rho_i$  be the smallest relative degree with respect to any input  $u_i$ , in the nominal system. This means that

$$y_i^{\rho_i} = L_f^{\rho_i} h_i(z) + L_g L_f^{\rho_i-1} h_i(z) u \quad (3.70)$$

where  $L_g L_f^{\rho_i-1} h_i(z) \in \mathbb{R}^m$ . That is, the row vector given by the Eq. (3.70) is nonzero (i.e., has at least a nonzero element). The existence of strong relative degree for each output such make possible combine these  $m$  equations,

$$\begin{bmatrix} y_1^{\rho_1} \\ y_2^{\rho_2} \\ \vdots \\ y_m^{\rho_m} \end{bmatrix} = A(z) + B(z)u$$

where  $A(z) \in \mathbb{R}^m$  and  $B(z) \in \mathbb{R}^{m \times m}$ . The nonlinear state feedback control law that provides input-output linearization of the nominal system can be expressed as

$$u = B(z)^{-1} (-A(z) + v(t)), \quad (3.71)$$

where  $v(t) \in \mathbb{R}^m$  is an auxiliary input. The nominal system has a **strong vector relative degree**  $(\rho_1, \dots, \rho_m)$  if the individual output  $y_i$  has strong relative degree  $\rho_i$  and  $B(z)$  is invertible. Considering the tracking problem, it is desired that  $y_i(t) \rightarrow y_{d_i}(t)$ , where  $y_{d_i}(t)$  is the smooth (at least  $\rho_i$ -times differentiable) desired trajectory. In order to design the control structure we define the error dynamics as  $e_i(t) = y_i(t) - y_{d_i}(t)$ . Therefore

$$\begin{bmatrix} e_1^{(\rho_1)}(t) \\ \vdots \\ e_m^{(\rho_m)}(t) \end{bmatrix} = \begin{bmatrix} y_1^{(\rho_1)}(t) - y_{d_1}^{(\rho_1)}(t) \\ \vdots \\ y_m^{(\rho_m)}(t) - y_{d_m}^{(\rho_m)}(t) \end{bmatrix}$$

In this case, it is possible to prove that  $\rho_1 = \dots = \rho_m = 1$ , therefore

$$\dot{e}_i(t) = L_f h_i(z) + L_g h_i(z) u + L_{\Delta f} h_i(z) + L_{\Delta g} h_i(z) u - \dot{y}_{d_i}(t)$$

Consequently, by applying the control law given by the Eq. (3.71), we have an uncertain dynamic system as follows:

$$\dot{e}(t) = \begin{bmatrix} \dot{e}_1(t) \\ \vdots \\ \dot{e}_m(t) \end{bmatrix} = \tilde{F}(z) + (I + \tilde{B}(z)) v - \dot{y}_d(t), \quad (3.72)$$

where

$$\tilde{B}(z) = \begin{bmatrix} L_{\Delta g_1} h_1(z) & \dots & L_{\Delta g_m} h_1(z) \\ \vdots & \ddots & \vdots \\ L_{\Delta g_1} h_m(z) & \dots & L_{\Delta g_m} h_m(z) \end{bmatrix} \begin{bmatrix} L_{g_1} h_1(z) & \dots & L_{g_m} h_1(z) \\ \vdots & \ddots & \vdots \\ L_{g_1} h_m(z) & \dots & L_{g_m} h_m(z) \end{bmatrix}^{-1}$$

$$\tilde{F}(z) = \begin{bmatrix} L_{\Delta f} h_1(z) \\ \vdots \\ L_{\Delta f} h_m(z) \end{bmatrix} - \tilde{B}(z) \begin{bmatrix} L_f h_1(z) \\ \vdots \\ L_f h_m(z) \end{bmatrix},$$

where  $I \in \mathbb{R}^{m \times m}$  is the identity matrix.

Taking

$$v = \dot{y}_d(t) - \frac{1}{\varepsilon} e(t) - a e(t) \quad (3.73)$$

the system becomes:

$$\dot{e}(t) = \tilde{F}(z) + \left( I + \tilde{B}(z) \right) \left[ \dot{y}_d(t) - \frac{1}{\varepsilon} e(t) - a e(t) \right] - \dot{y}_d$$

or equivalently

$$\dot{e}(t) = - \left( \frac{1}{\varepsilon} + a \right) \left( I + \tilde{B}(z) \right) e(t) + \left( \tilde{F}(z) + \tilde{B}(z) \dot{y}_d \right) \quad (3.74)$$

In order to achieve the Robust Practical Synchronization a high gain mode control law is designed such that trajectories of the closed-loop system are attracted to a ball of radius  $\varepsilon$  and with center at the origin.

**Proposition 3.1.** If  $\|f_i(z)\| > \|\Delta f_i(z)\|$  and  $\|g_{i,j}(z)\| > \|\Delta g_{i,j}(z)\|$  given by Eq. 3.67, for all  $i \in \{1, \dots, n\}$  and  $j \in \{1, \dots, m\}$ , the dynamic system given by the Eq. (3.74) converges asymptotically to a ball of radius  $\varepsilon$  and with center at the origin.  $\diamond$

Notice that although the conditions  $\|f_i(z)\| > \|\Delta f_i(z)\|$  and  $\|g_{i,j}(z)\| > \|\Delta g_{i,j}(z)\|$  are restrictive in general terms, from a practical point of view, if such conditions are not fulfilled implies that we do not have a sufficiently accurate model of the system. Moreover, note that the previous proposition automatically fulfills the requirements for practical synchronization according to its definitions. Thus, the control law guarantees the robust synchronization of the master and slave systems and therefore ensures that the the real systems follow the optimal batch operation model. In order to illustrate the important issues we display in the following section the numerical simulation performed using the method here developed.

## Numerical Simulations

In order to apply the aforementioned scheme, the MIMO Fed-batch Reactor studied in the section 3.5 and given by the Eqs. (3.48)-(3.50) is rewritten as follows.

$$f(Z) = \begin{bmatrix} Z_1 \mu(Z_2) \\ -\frac{Z_1}{Y} \mu(Z_2) \end{bmatrix}, \quad g_1(Z) = \begin{bmatrix} -Z_1 \\ S_f - Z_2 \end{bmatrix}, \quad g_2(Z) = \begin{bmatrix} -Z_1 \\ -Z_2 \end{bmatrix}$$

$$\Delta f(Z) = \begin{bmatrix} -Z_1 r_d \\ 0 \end{bmatrix}, \quad \Delta g_1(Z) = \begin{bmatrix} 0 \\ \delta \end{bmatrix}, \quad \Delta g_2(Z) = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

$$h_1(z) = z_1, \quad h_2(z) = z_2. \quad (3.75)$$

According to the proposed approach, it is possible to show that the dynamic error is given by the following equations

$$\begin{bmatrix} \dot{e}_1(t) \\ \dot{e}_2(t) \end{bmatrix} = \begin{bmatrix} -r_d z_1 + w_1 \\ \frac{\delta z_2 \mu_{\max}(z_2 Y + z_1)}{S_f (K_2 z_2^2 + z_2 + K_1) Y} - \frac{\delta z_2 w_1}{z_1 S_f} + \left(1 + \frac{\delta}{S_f}\right) w_2 \end{bmatrix} - \begin{bmatrix} \dot{y}_{d_1}(t) \\ \dot{y}_{d_2}(t) \end{bmatrix}. \quad (3.76)$$

In order to examine the robustness of the approach, various perturbation were imposed to the bioreactor that are presented graphically in the Figure 3.3.

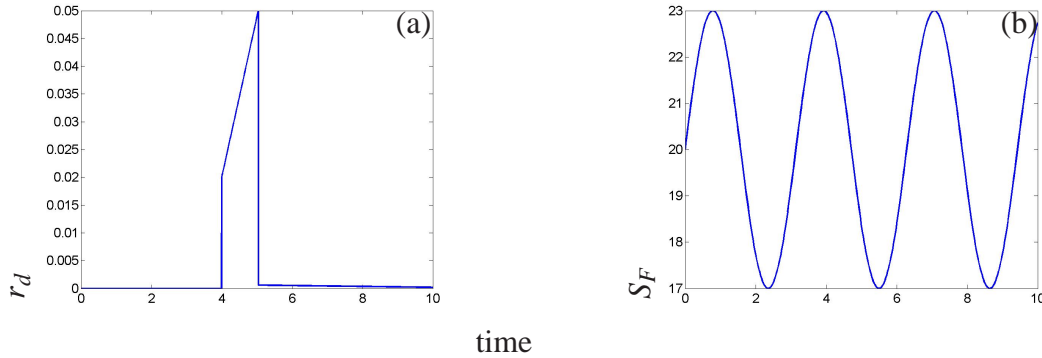


Figure 3.3: The numerical simulation – blue solid lines represent the perturbation imposed over the bioreactor. Plot (a) represents the evolution of dead cell rate ( $r_d$ ) in time and plot (b) the variation of fed substrate concentration ( $S_F$ ) in time.

Figures 3.4-3.5 shows the performance of the scheme proposed in this chapter under the conditions mentioned before. As one can see the scheme does a very good job.

### Sliding Control Law

In this section, another alternative to achieve the practical synchronization is presented. This alternative is the sliding mode control, which is characterized by a control law containing the discontinuous nonlinearity  $\text{sgn}(\cdot)$ . This discontinuity causes some theoretical as well as practical problems (the chattering). The approach used in this section is based on the work of Behtash [47]. In the following figures, we present the numerical results of the application of the sliding control to the same problem treated in previously.

To achieve the practical synchronization the system given by the Eq. (3.72) should converge to a neighborhood of zero. According to Behtash [47], using the control structure it is possible to achieve the task.

$$v = \dot{y}_d(t) - \text{sgn}(e(t)) - a e(t) \quad (3.77)$$

From the graphics given in the Fig. 3.6-3.7, it is possible to conclude that although the sliding controller achieves the synchronization the presence of chattering could be very damaging for the actuators. On the other hand, the high gain controllers could be quite sensitive to noise.

### Sliding Control Law and Sliding Observer

Taking into account not only the robustness against noise, but also in the sense of state estimation, in this section a sliding control based on sliding state observer is proposed. The approach used in this section is based on the work of Wang et al.[46].

From the Fig. 3.8, it is possible to conclude that although the general performance in this case is bad in comparison with the two previous cases, in this case is much more robust in the sense that we need less outputs, i.e, we estimate one of the states of the system.

## 3.7 Concluding Remarks

In this chapter, the problem of following an optimal batch operation model for a Bioreactor in the presence of uncertainties was tackled through a master-slave synchronization approach. The advantages of this idea are that we can easily perform the optimization for a simplified model and then design the controller such that the real bioreactor is synchronized with the optimized one in spite of unknown dynamics and perturbations.

This general method can also be used for transition from batch and fed-batch to continuous culture or for repeated fed-batch operations. It is possible also to note that our scheme is very general allowing the possibility of following the optimal batch operation model under any unknown dynamics. Moreover, this method represents a new application field for many of the system synchronization results.

In addition, the reachability of a simple nonlinear fed-batch fermentation process model and the exact feedback linearization are investigated in this chapter. It is shown that the inclusion of the extra input variable guarantees the rank of the reachability distribution and make plausible the usage of the exact linearization technique. This type of results complement the work developed by Szerderkényi et al.[39] in the understanding of the controllability properties of the fed-batch fermentation process.

## 3.8 Notation and Definitions

In this section, we briefly recall some basic concepts from differential geometry.

**Definition**[Lie derivative]: This operation involves a real-valued function  $\Lambda$  and a vector field  $\zeta$ , both defined in a subset  $U \in \mathbb{R}^n$ . From these, a new smooth real-valued function is defined, denoted  $L_\zeta \Lambda(X)$ , and defined as

$$L_\zeta \Lambda(X) = \frac{\partial \Lambda}{\partial X} \zeta(X) = \sum_{i=1}^n \frac{\partial \Lambda}{\partial X_i} \xi_i(X).$$

The repeated use of this operation is possible. Thus, it is possible to construct the  $j$ th Lie derivative. This can be expressed as  $L_\zeta^j(\Lambda)(X(t))$  and is defined inductively by

$$\begin{aligned} L_\zeta^0 \Lambda(X) &= \Lambda(X) \\ L_\zeta^j \Lambda(X) &= \frac{\partial}{\partial X} (L_\zeta^{j-1} \Lambda(X)) \zeta(X). \end{aligned}$$

□

**Proposition**[Lie bracket]: This operation involves two vector fields  $\zeta$  and  $\xi$ , both defined in a subset  $U \in \mathbb{R}^n$ . From these ones, a new smooth vector field is constructed, denoted  $[\zeta, \xi](X)$  and defined as

$$[\zeta, \xi] = \frac{\partial \xi}{\partial X} \zeta(X) - \frac{\partial \zeta}{\partial X} \xi(X)$$

where  $\frac{\partial \xi}{\partial X}$  and  $\frac{\partial \zeta}{\partial X}$  stand for the jacobian matrices of  $\xi$  and  $\zeta$ , respectively. Repeated bracketing of vector field  $\xi$  with the same vector field  $\zeta$  is possible. Whenever this is required, in order to avoid a notation of the form  $[\zeta, [\zeta, \dots, [\zeta, \xi]]]$ , we define such operation recursively, as

$$\text{ad}_\zeta^k \xi(X) = [\zeta, \text{ad}_\zeta^{k-1} \xi(X)].$$

for any  $k \geq 1$ , setting  $\text{ad}_\zeta^0 \xi(X) = \xi(X)$ . □



This operation is characterized by three basic properties that are summarized in the following statement.

**Proposition:** The Lie Bracket of a vector field has the following properties:

i) it is bilinear over  $\mathbb{R}$ , i.e., if  $\zeta_1, \zeta_2, \xi_1, \xi_2$  are vectors fields and  $r_1, r_2$  real numbers, then

$$\begin{aligned} [r_1 \zeta_1 + r_2 \zeta_2, \xi_1] &= r_1 [\zeta_1, \xi_1] + r_2 [\zeta_2, \xi_1] \\ [\zeta_1, r_1 \xi_1 + r_2 \xi_2] &= r_1 [\zeta_1, \xi_1] + r_2 [\zeta_1, \xi_2] \end{aligned}$$

ii) it is skew commutative, i.e.

$$[\zeta_1, \xi_1] = -[\xi_1, \zeta_1]$$

iii) it satisfies the Jacobi identity

◇

### 3.9 Proof of proposition 3.1

Consider the following Lyapunov function

$$V(e) = \frac{1}{2} e^T I e \Rightarrow \dot{V}(e) = e^T I \dot{e} \quad (3.78)$$

$$(3.79)$$

$$\dot{V}(e) = e^T I \left( - \left( \frac{1}{\varepsilon} + a \right) \left( I + \tilde{B}(z) \right) e + \left( \tilde{F}(z) + \tilde{B}(z) \dot{y}_d \right) \right)$$

$$\dot{V}(e) = - \left( \frac{1}{\varepsilon} + a \right) e^T e - \left( \frac{1}{\varepsilon} + a \right) e^T \tilde{B}(z) e + e^T \left( \tilde{F}(z) + \tilde{B}(z) \dot{y}_d \right).$$

Let

$$\varepsilon_1 = \left\| \tilde{F}(z) \right\|_{\infty}; \quad \varepsilon_2 = \left\| \tilde{B} \right\|_{\infty}; \quad d = \left\| \dot{y}_d \right\|_{\infty}$$

Note that

$$|e^T e| \leq |e|_1^2; \quad \left\| e^T \left( \tilde{F}(z) + \tilde{B}(z) \dot{y}_d \right) \right\|_{\infty} \leq (\varepsilon_1 + \varepsilon_2 d) |e|$$

when it is assumed that

$$\left| e^T \tilde{B}(z) e \right| \leq \left\| \tilde{B}(z) \right\| |e|_1^2. \quad (3.80)$$

Therefore

$$\dot{V}(e) \leq -\left(\frac{1}{\varepsilon} + a\right) |e|_1^2 - \left(\frac{1}{\varepsilon} + a\right) \left\| \tilde{B}(z) \right\| |e|_1^2 + (\varepsilon_1 + \varepsilon_2 d) |e|_1$$

$$\dot{V}(e) \leq -(1 + \varepsilon a)(1 + \varepsilon_2) |e|_1^2 + \varepsilon (\varepsilon_1 + \varepsilon_2 d) |e|_1$$

Let

$$c_1 = (1 + \varepsilon a)(1 + \varepsilon_2); \quad c_2 = (\varepsilon_1 + \varepsilon_2 d) \quad (3.81)$$

then

$$\dot{V}(e) \leq -c_1 |e|_1^2 + \varepsilon c_2 |e|_1$$

$$\dot{V}_i(e) \leq -c_1 |e_i|_1^2 + \varepsilon c_2 |e_i|_1.$$

Therefore

$$(\varepsilon \rightarrow 0) \Rightarrow (\dot{V}(e) < 0).$$

This complete the proof.  $\square$

**Proposition** The inequality given by the Eq. 3.80 stand in our case.

$$\left| e^T \tilde{B}(z) e \right| \leq \left\| \tilde{B}(z) \right\|_{\infty} |e|_1^2$$

◇

Note that

$$\left| e^T \tilde{B}(z) e \right| = \left| \sum_{i=1}^n \sum_{j=1}^n [\tilde{B}(z)]_{i,j} e_i e_j \right|$$

$$|e|_1^2 = \left( \sum_{j=1}^n |e_j| \right)^2 = \sum_{i=1}^n \sum_{j=1}^n \alpha_{i,j} |e_i| |e_j|$$

where

$$\alpha_{i,j} \begin{cases} i = j, \alpha_{i,j} = 1 \\ i \neq j, \alpha_{i,j} = 2 \end{cases}$$

$$\|\tilde{B}(z)\|_{\infty} = \max \left( \sum_{j=1}^n [\tilde{B}(z)]_{1,j}, \dots, \sum_{j=1}^n [\tilde{B}(z)]_{m,j} \right).$$

On the other hand

$$|e^T \tilde{B}(z) e| \leq \sum_{i=1}^n \sum_{j=1}^n |[\tilde{B}(z)]_{i,j}| |e_i| |e_j|.$$

Let

$$\sum_{k=1}^n |[\tilde{B}(z)]_{*,k}| = \|\tilde{B}(z)\|_{\infty}.$$

Then

$$\|\tilde{B}(z)\|_{\infty} |e|_1^2 = \left( \sum_{k=1}^n |[\tilde{B}(z)]_{*,k}| \right) \left( \sum_{i=1}^n \sum_{j=1}^n \alpha_{i,j} |e_i| |e_j| \right),$$

$$\|\tilde{B}(z)\|_{\infty} |e|_1^2 = \sum_{i=1}^n \sum_{j=1}^n \sum_{k=1}^n \alpha_{i,j} |[\tilde{B}(z)]_{*,k}| |e_i| |e_j|.$$

where we can note that

$$\sum_{i=1}^n \sum_{j=1}^n |[\tilde{B}(z)]_{i,j}| |e_i| |e_j| \leq \sum_{i=1}^n \sum_{j=1}^n \sum_{k=1}^n \alpha_{i,j} |[\tilde{B}(z)]_{*,k}| |e_i| |e_j|.$$

Therefore

$$|e^T \tilde{B}(z) e| \leq \|\tilde{B}(z)\|_{\infty} |e|_1^2$$

This completes the proof.  $\square$

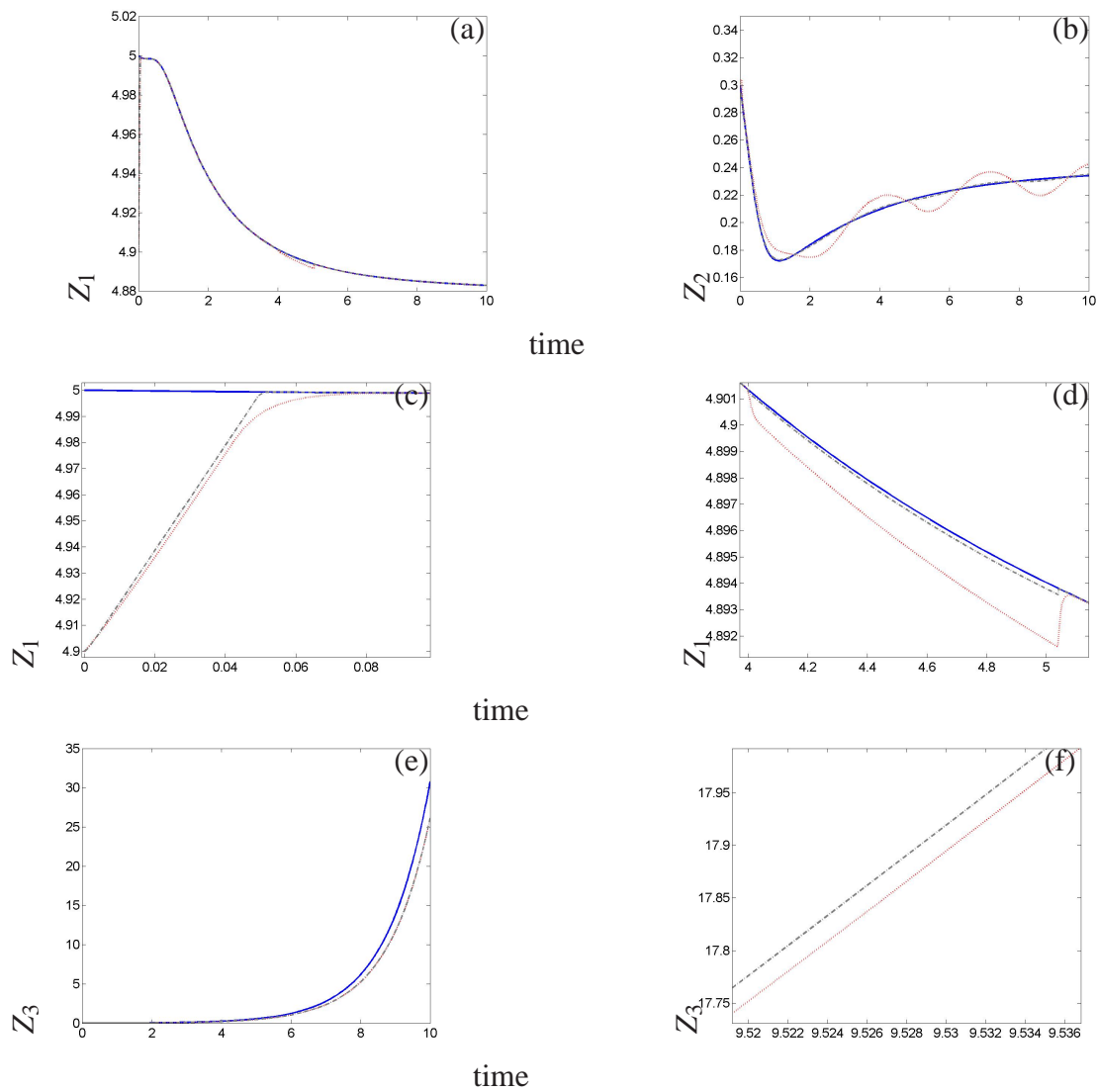


Figure 3.4: The numerical simulation – blue solid lines represent the master system, red dotted lines represent slave with  $\epsilon = 0.01$  and grey dash-dotted lines represent slave with  $\epsilon = 0.001$ . Plot (a) represents the evolution of biomass concentration in time and plot (b) the variation of substrate concentration in time. Plot (c) represents a detail of the initial evolution of biomass concentration in time and plot (d) a detail of the evolution of biomass concentration when the perturbation over the dead cell rate occurs. Plot (e) represents evolution of bioreactor's volume in time and plot (f) the difference in the evolution between the control laws with  $\epsilon = 0.01$  and  $\epsilon = 0.001$ , respectively.

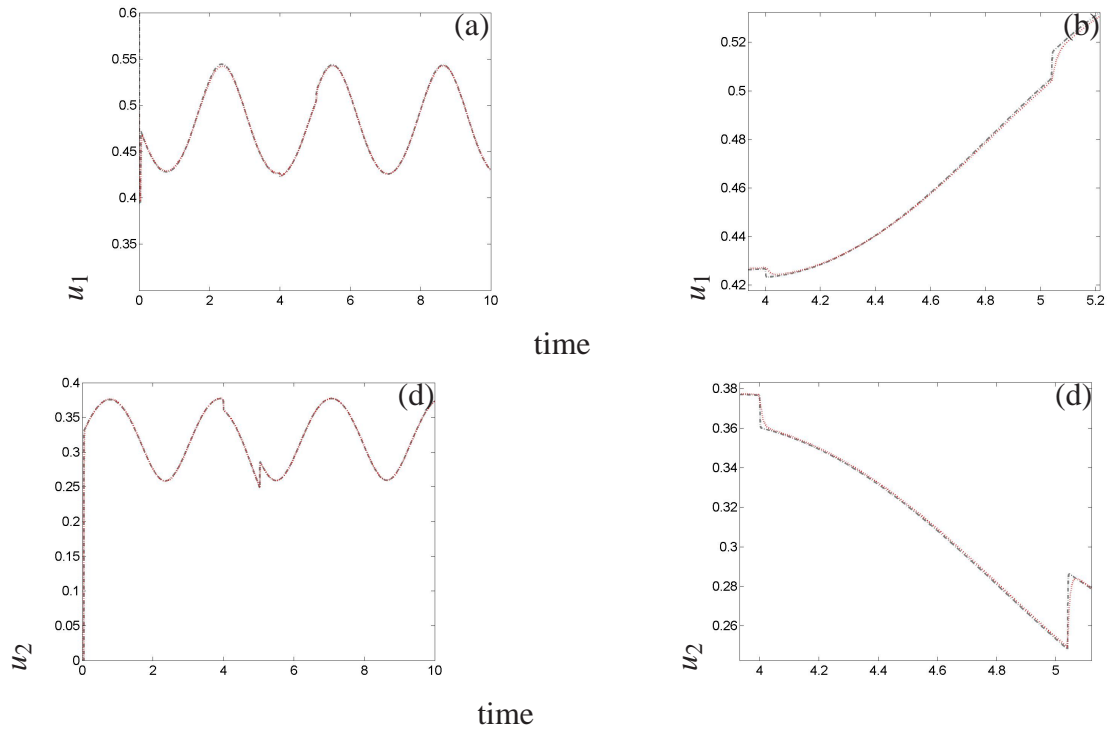


Figure 3.5: The numerical simulation – blue solid lines represent the master system, red dotted lines represent slave with  $\varepsilon = 0.01$  and grey dash-dotted lines represent slave with  $\varepsilon = 0.001$ . Plot (a) represents the evolution of input  $u_1$  in time and plot (b) a detail of the evolution of input  $u_1$  when the perturbation over the dead cell rate occurs. Plot (c) represents the evolution of input  $u_2$  in time and plot (d) a detail of the evolution of input  $u_2$  when the perturbation over the dead cell rate occurs.

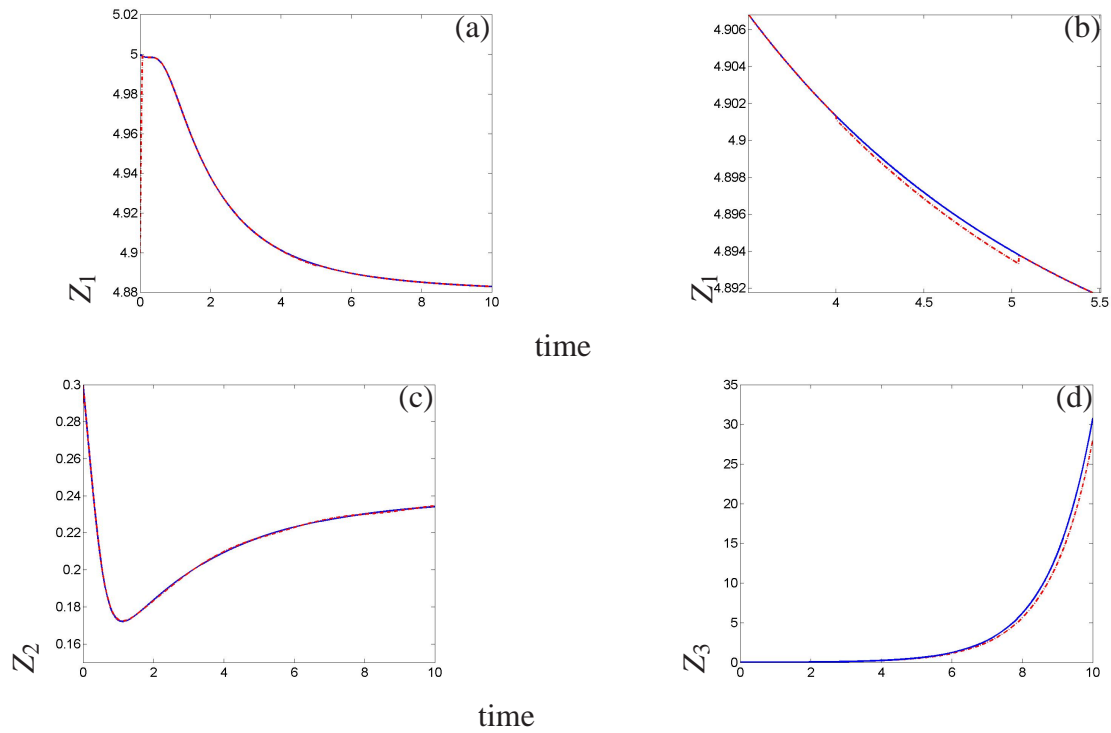


Figure 3.6: The numerical simulation – blue solid lines represent the master system and red dash-dotted lines represent Slave System under the sliding control law. Plot (a) represents the evolution of biomass concentration in time and plot (b) a detail of the evolution of biomass concentration when the perturbation over the dead cell rate occurs. Plot (c) represents the evolution of substrate concentration in time and plot (d) represents evolution of bioreactor's volume in time.

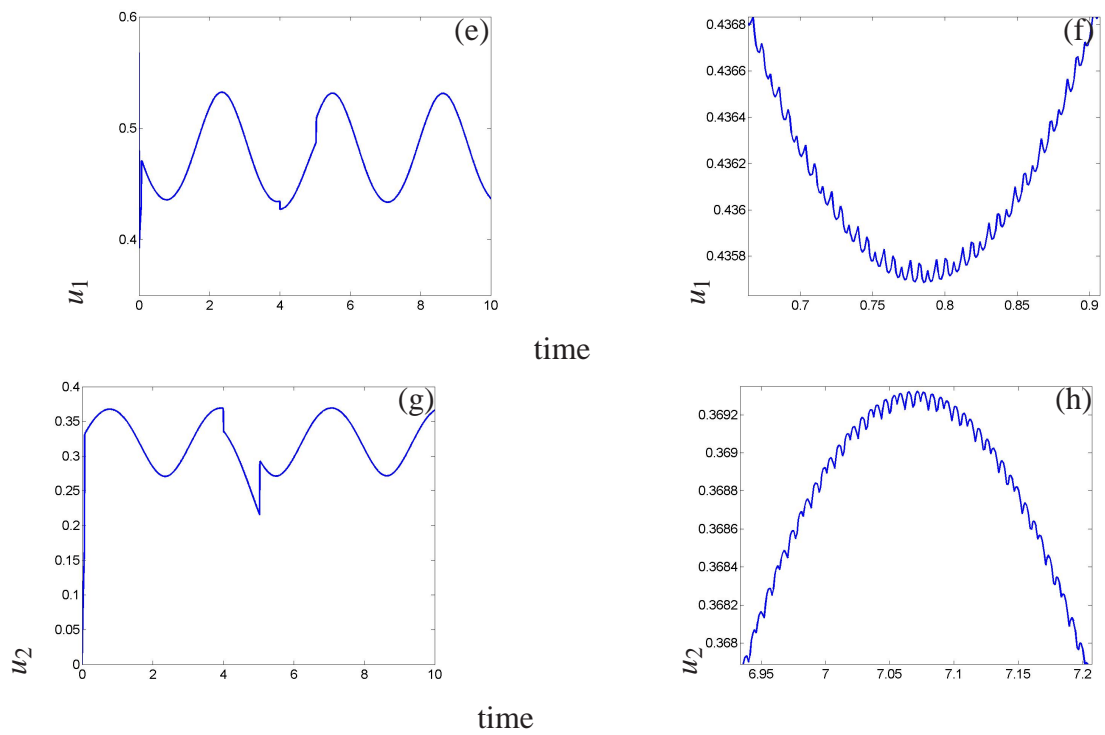


Figure 3.7: The numerical simulation – blue solid lines represent the master system and red dash-dotted lines represent slave system under the sliding control law. Plot (e) represents the evolution of input  $u_1$  in time and plot (f) a detail of the evolution of input  $u_1$  where it is possible to appreciate the chattering effect.

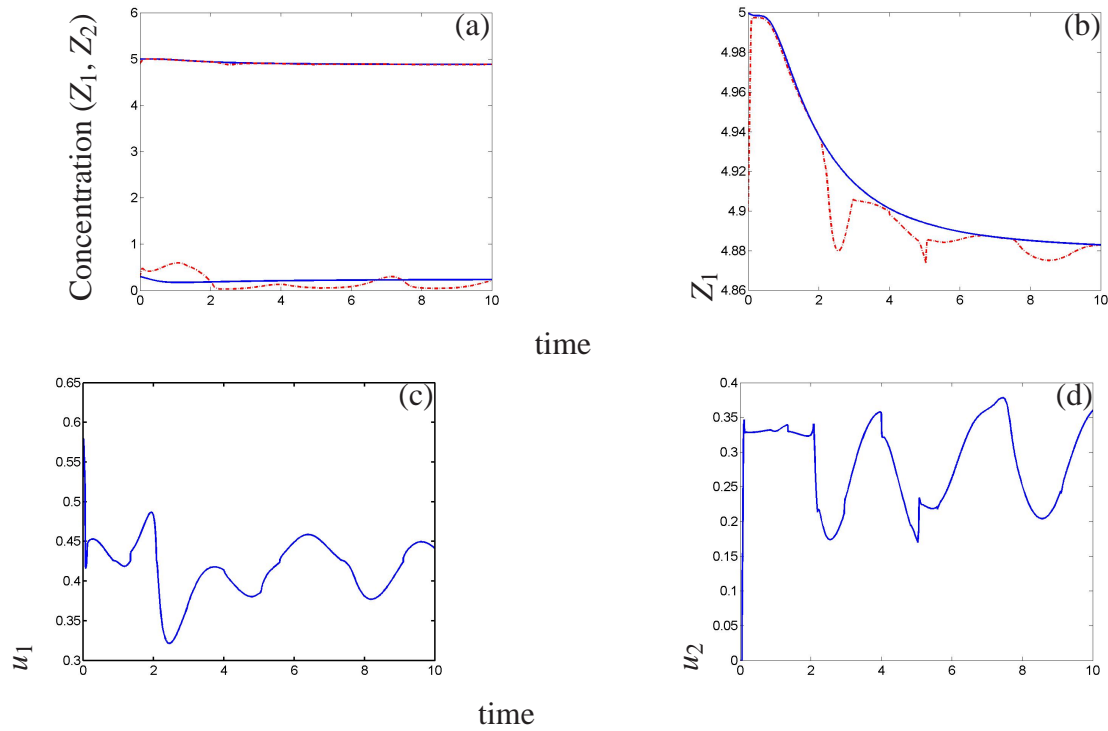


Figure 3.8: The numerical simulation – blue solid lines represent the master system and red dash-dotted lines represent slave system under the sliding control law and coupled with a sliding observer. Plot (a) represents the evolution of substrate and biomass concentration in time and plot (b) a detail of the evolution of biomass concentration. Plot (c) represents the evolution of input  $u_1$  concentration in time and plot (d) a detail of the evolution of input  $u_2$ .



4

---

## Inferring Mixed Culture Growth

---

### 4.1 Inferring mixed-culture growth from total biomass data in a wavelet approach

This work was performed in collaboration with the Ph.D. student Pilar Escalante-Minakata, Dr. Jose Salome Murguía Ibarra and Professor Dr. Haret C. Rosu [48]. Here, we shown that the presence of mixed-culture growth in batch fermentation processes can be very accurately inferred from total biomass data by means of the wavelet analysis for singularity detection. This is accomplished by considering simple phenomenological models for the mixed growth and the more complicated case of mixed growth on a mixture of substrates. The main quantity provided by the wavelet analysis is the Hölder exponent of the singularity that we determine for our illustrative examples. The numerical results point to the possibility that Hölder exponents can be used to characterize the nature of the mixed-culture growth in batch fermentation processes with potential industrial applications. Moreover, the analysis of the same data affected by the common additive Gaussian noise still lead to the wavelet detection of the singularities although the Hölder exponent is no longer a useful parameter.

### 4.2 Introduction

The growth of microbial species in media containing two or several growth-limiting substrates is of great importance in biotechnology and bioengineering. The mixed-culture growth occurs in many industrial processes. A first significant class of such processes is the traditional fermented foods and beverages in which either endemic microorganisms or an inoculum with selected microorganisms are used, see for instance [49]. Some beverages get two or more different microorganisms in the inoculum with the purpose to provide a desired flavor. Evidence of

this influence are presented in the recent paper of [50], in which the role of different yeast interactions on the wine flavor is discussed. However, the phenomenological details and the theory of the time evolution of the fermentation are as yet poorly understood. We can also mention the interesting case of the bioethanol production, in which the substrates used for fermentation typically consist of a mixture of glucose and fructose. Bioethanol is the product obtained from the metabolism of microbe mixtures feeding with this combination of hexoses and pentoses, see e.g., [51]. The last relevant example we give is bioremediation, in which gasoline and chemical spills generally yield a complex mixture of water-soluble organic compounds. In gasoline spills, for instance, the four basic compounds are benzene, toluene, ethylbenzene, and xylene. The consumption of this mixture by microorganisms is what is defined as the bioremediation process.

In all the aforementioned cases, the presence of different populations of microorganisms and substrates is a key factor in the quality and quantity of the final product. Therefore, it is quite useful to detect the presence or lacking of process of mixed-culture growth type. Their presence could be used as an estimate of the right evolution of the process in its early stage. In addition, a rapid and reasonably accurate test is always useful for saving time and helping to take quick decisions. It is quite clear then that the biomass concentration is one of the most needed quantity that should be measured in fermentation monitoring. The most popular method to get the biomass concentration is by means of the measurement of the optical density of centrifugalized samples. However, this procedure has limited usefulness because it cannot distinguish neither the living cells from the dead ones, nor the different types of microorganisms involved in the process. In some cases it is also possible to correlate the total biomass concentration with the values of the redox potential of the fermentation.

Recently, new techniques have emerged to quantify the biomass and distinguish the different microorganisms present in a mixed-culture. Some of them based on sophisticated equipment ([53], [54] and [55]) and others resides on molecular biology techniques ([51] and [52]). All these techniques are very promising in the study of the dynamics of the mixed-culture growth, although, they require expensive or complicated procedures. In this chapter, we show that it is possible to infer mixed-culture growth of microorganisms from their total biomass data, without using such complicated techniques. The alternative procedure that we put forth here is based on treating the total biomass data by means of the wavelet approach for detection of singularities in the growth curves. The idea is to treat the mixed growth curves as more or less regular signals that can nevertheless display singularities due to their compound structure. In the wavelet literature there exist fundamental papers in which it has been shown that the wavelet techniques are very efficient in detecting any type of singularities.

The rest of the chapter is organized as follows. In Section 4.3 and 4.4, we introduce the dynamics of the mixed-growth type and discuss its basic assumptions. Next, in Section 4.5, the method of the wavelet singularity analysis is briefly presented, whereas its application to

the mixed type dynamical curves is enclosed in Section 4.6. A conclusion section ends up the chapter. An appendix containing the standard definitions of Hölder exponents of singularities of functions is included as well.

### 4.3 A simple mixed-growth model

The technology of batch processes is well developed and numerous products are obtained in this way. Some products such as food, beverages, and pharmaceutical ones require precise tracking of the batch information for safety and regulatory purposes. The primary objective of monitoring batch processes is to ensure that significant and sustained changes in the quality of the product (caused by disturbances and/or faults) are detected as soon as possible. In that sense, the rapid detection of singularities in the output of the batch processes offers an interesting solution. The wavelet analysis for singularity detection is by now well established but there was no direct application to infer mixed-growth in the case of batch biochemical processes.

In order to achieve this task we will consider here a fermentation process consisting of a perfectly stirred tank, where no streams are fed into it. In the batch fermenter the substrate is converted by biomass into additional biomass and products. The general unstructured mass balances for the well-mixed bioreactor can be represented by the following equations for the concentrations of the cells and substrates:

$$\frac{dx_{1,i}}{dt} = x_{1,i} \mu_i(x_{2,i}) \quad (4.1)$$

$$\frac{dx_{2,i}}{dt} = -\frac{x_{1,i}}{Y_i} \mu_i(x_{2,i}) \quad (4.2)$$

where  $x_{1,i}$  represent the biomass concentrations,  $x_{2,i}$  substrate concentrations and  $Y_i$  is the biomass yield,  $\mu_i(x_{2,i})$  is the specific growth rate and  $i \in \mathbb{Z}^+$  represent the  $i$ -th species, allowing for the possibility of multiple kinds of substrates and microorganisms. The growth rate relates the change in biomass concentrations to the substrate concentrations. Two types of relationships for  $\mu_i(x_{2,i})$  are commonly used: the substrate saturation model (Monod Equation) and the substrate inhibition model (Haldane Equation). Both cases will be treated here. The substrate inhibited growth can be described by

$$\mu_i(x_{2,i}) = \frac{\mu_{max_i} x_{2,i}}{K_{1_i} + x_{2,i} + K_{2_i} x_{2,i}^2} \quad (4.3)$$

where  $K_{1_i}$  is the saturation (or Monod) constant,  $K_{2_i}$  is the inhibition constant and  $\mu_{max_i}$  is the maximum specific growth rate. The value of  $K_{1_i}$  expresses the affinity of biomass for substrate. The Monod growth kinetics can be considered as a special case of the substrate inhibition kinetics with  $K_{2_i} = 0$  when the inhibition term vanishes. For the sake of simplicity, we will consider only two species and two substrates. Moreover, we consider that it is possible

to measure only the total biomass concentration. That means that the output of the system ( $y$ ) will be given by

$$y = \sum_{i=1}^m x_{1,i} \quad (4.4)$$

where  $m$  is the number of species of microorganisms growing in the bioreactor (in this work  $m = 2$ ). We focus on the following four cases:

- I The microorganism and substrate concentrations have the same initial conditions, but different growth rates, one with a Haldane type and one with a Monod type. In addition, quite different values of the Monod constant will be taken into account.
- II The microorganism and substrate concentrations have different initial conditions, but the same growth rates.
- III The microorganism and substrate concentrations have different initial conditions and different growth rates, one with a Haldane type and one with a Monod type.
- IV The microorganism and substrate concentrations have the same initial conditions and the same growth rates, but with different values of the maximal growth rate.

Table 4.1: The initial conditions and the values of the employed parameters of the mixed-growth process model.

Symbol	Meaning	Values				Units
		Case I	Case II	Case III	Case IV	
$\mu_{max_1}$	Maximal growth rate	1	1	1	0.9	[l/h]
$K_{1_1}$	Saturation parameter	0.03	0.03	0.03	0.03	[g/l]
$K_{2_1}$	Saturation parameter	0.5	0.5	0.02	0.5	[g/l]
$Y_1$	Yield coefficient	0.5	0.5	0.5	0.5	—
$x_{1_1}^0$	Initial biomass conc.	0.1	0.1	0.1	0.25	[g/l]
$x_{2_1}^0$	Initial substrate conc.	10	10	10	10	[g/l]
$\mu_{max_2}$	Maximal growth rate	1	1	1	1	[l/h]
$K_{1_2}$	Saturation parameter	0.3	0.03	0.03	0.03	[g/l]
$K_{2_2}$	Inhibition parameter	0	0.5	0	0.5	[l/g]
$Y_2$	Yield coefficient	0.5	0.5	0.5	0.5	—
$x_{1_2}^0$	Initial biomass conc.	0.1	0.2	0.1	0.25	[g/l]
$x_{2_2}^0$	Initial substrate conc.	10	5	6	10	[g/l]

Table 4.1 shows the variables and parameter values used to simulated the two species growing in the two different substrates, under the four cases under consideration.

## 4.4 Mixed cultures on mixtures of substrates

When microbes are grown in a batch reactor containing a surplus of two substrates, one of the substrates is generally exhausted before the other, leading to the appearance of two successive exponential growth phases. This phenomenon could be noticeable at simple view in the biomass signal or unnoticed due to its nature or due to additive noise present in the signal. In general, such type of phenomenon is known as growth of mixed cultures on mixtures of substrates (MCMS).

The growth of MCMS is a phenomenon of practical and theoretical interest. The fundamental understanding of this problem has impact on many practical fields such as food processing, production of ethanol from renewable resources, bioremediation and microbial ecology, among many others. To study the usage of the wavelet approach in the detection of MCMS growth, we consider the recent model proposed by Reeves 2004 [56], which takes into account such type of growth.

Within this section, the index  $i$  will denote the species number, and the index  $j$  will stand for the substrate number. Thus,  $c_i$  denotes the concentration of the  $i$ th species,  $s_j$  denotes the concentration of the  $j$ th substrate,  $e_{ij}$  denotes the concentration of the lumped system of inducible enzymes catalyzing the uptake and peripheral catabolism of  $s_j$  by  $c_i$ . Here,  $c_i$  and  $s_j$  are based on the volume of the chemostat, and expressed in the units gdw/l and g/l, respectively.  $e_{ij}$  is based on the dry weight of the biomass, and expressed in the units g/gdw.

$$r_{ij}^s = V_{ij}^s e_{ij} \frac{s_j}{K_{ij}^s + s_j} \quad (4.5)$$

$$r_{ij}^x = k_{ij}^x x_{ij} \quad (4.6)$$

$$r_{ij}^e = V_{ij}^e \frac{x_{ij}}{K_{ij}^e + x_{ij}} \quad (4.7)$$

$$r_{ij}^{ast} = k_{ij}^{ast} \quad (4.8)$$

$$r_{ij}^d = K_{ij}^d e_{ij} \quad (4.9)$$

$$\frac{ds_j}{dt} = D(s_j^f - s_j) - r_{1j}^s c_1 - r_{2j}^s c_2 \quad (4.10)$$

$$\frac{de_{ij}}{dt} = V_{ij}^e \frac{e_{ij} \sigma_{ij}}{K_{ij}^e + e_{ij} \sigma_{ij}} + k_{ij}^{ast} - k_{ij}^d e_{ij} - r_i^g e_{ij} \quad (4.11)$$

$$\frac{dc_i}{dt} = (r_i^g - D) c_i, \quad (4.12)$$

where

$$\bar{K}_{ij}^e = \frac{K_{ij}^e k_{ij}^x}{V_{ij}^s}, \quad \sigma_{ij} = \frac{s_j}{K_{ij}^s + s_j} \quad (4.13)$$

$$r_i^s = Y_{i1} r_{i1}^s + Y_{i2} r_{i2}^s \quad (4.14)$$

Reeves et al [56] comment that a plausible experimental situation is the case of *Escherichia coli* and *Pseudomonas aeruginosa*, in which, *E. coli* prefers a sugar over an organic acid, and *P. aeruginosa* prefers the organic acid over the sugar.

Table 4.2: Parameter values used in the MCMS growth model [56]

$V_{11}^s = 1000$	$V_{12}^s = 1000$	$V_{21}^s = 1000$	$V_{22}^s = 1000$	g/g h
$K_{11}^s = 0.01$	$K_{12}^s = 0.01$	$K_{21}^s = 0.01$	$K_{22}^s = 0.01$	g/l
$V_{11}^e = 0.0025$	$V_{12}^e = 0.0020$	$V_{21}^e = 0.0006$	$V_{22}^e = 0.0036$	g/gdw h
$\bar{K}_{11}^e = 0.0017$	$\bar{K}_{12}^e = 0.0032$	$\bar{K}_{21}^e = 0.0013$	$\bar{K}_{22}^e = 0.0030$	g/gdw
$k_{11}^d = 0.01$	$k_{12}^d = 0.01$	$k_{21}^d = 0.01$	$k_{22}^d = 0.01$	l/h
$k_{11}^* = 10^{-2} V_{11}$	$k_{12}^* = 10^{-2} V_{12}$	$k_{21}^* = 10^{-2} V_{21}$	$k_{22}^* = 10^{-2} V_{22}$	g/gdw h
$Y_{11} = 0.41$	$Y_{12} = 0.24$	$Y_{21} = 0.35$	$Y_{22} = 0.20$	g/g

In order to have the batch regime in the bioreactor we set parameter  $D = 0$  and also employ Reeves' parameters  $s_1^f = 1$  and  $s_2^f = 2$ . Table 4.2 shows the rest of the parameter values used to simulate the growth of the two species on the two different substrates in the MCMS conditions.

## 4.5 Measuring regularity with the wavelet transform

Let us think of the total biomass of the mixed-growth curves as a signal. In general, performing the analysis of a signal means to find the regions of its regular and singular behavior. Usually the singularities are very specific features for signal characterization. As it has been pointed in the seminal paper of [62], the regularity of a signal treated as a function can be characterized by Hölder exponents. The wavelet transform has been demonstrated to be a tool exceptionally well suited for the estimation of Hölder exponents (for their definitions see the Appendix).

### 4.5.1 The wavelet transform

Let  $L^2(\mathbb{R})$  denote the space of all square integrable functions on  $\mathbb{R}$ . In signal processing terminology,  $L^2(\mathbb{R})$  is the space of functions with finite energy. Let  $\psi(t) \in L^2(\mathbb{R})$  be a fixed function. The function  $\psi(t)$  is said to be a wavelet if and only if its Fourier transform,  $\hat{\psi}(\omega) = \int e^{i\omega t} \psi(t) dt$ , satisfies

$$C_\psi = \int_0^\infty \frac{|\hat{\psi}(\omega)|^2}{|\omega|} d\omega < \infty. \quad (4.15)$$

The non-divergent relation given by Eq. (4.15) is called the *admissibility condition* in wavelet theory, see for instance [60] and [62]. It implies that the wavelet must have a zero average on the real line

$$\int_{-\infty}^{\infty} \psi(t) dt = \hat{\psi}(0) = 0, \quad (4.16)$$

and therefore it must be oscillatory. In other words,  $\psi$  must be a sort of *wave* ([60, 62]). Based on  $\psi(t)$ , one defines the functions  $\psi_{a,b}$  as follows

$$\psi_{a,b}(t) = \frac{1}{\sqrt{a}} \psi\left(\frac{t-b}{a}\right), \quad (4.17)$$

where  $b \in \mathbb{R}$  is a translation parameter, while  $a \in \mathbb{R}^+$  ( $a \neq 0$ ) is a dilation or scale parameter. The factor  $a^{-1/2}$  is a normalization constant such that  $\psi_{a,b}$  has the same energy for all scales  $a$ . One notices that the scale parameter  $a$  in Eq. (4.17) is a measure of the dilations of the spatial variable  $(t-b)$ . In the same way the factor  $a^{-1/2}$  measures the dilations of the values taken by  $\psi$ . Because of this, one can decompose a square integrable function  $f(t)$  in terms of the dilated-translated wavelets  $\psi_{a,b}(t)$ . We define the wavelet transform (WT) of  $f(t) \in L^2(\mathbb{R})$  by

$$W_f(a,b) = \langle f, \psi_{a,b} \rangle = \int_{-\infty}^{\infty} f(t) \bar{\psi}_{a,b}(t) dt = \frac{1}{\sqrt{a}} \int_{-\infty}^{\infty} f(t) \bar{\psi}\left(\frac{t-b}{a}\right) dt, \quad (4.18)$$

where  $\langle \cdot, \cdot \rangle$  is the scalar product in  $L^2(\mathbb{R})$  defined as  $\langle f, g \rangle := \int f(t) \bar{g}(t) dt$ , and the bar symbol denotes complex conjugation. The WT given by Eq. (4.18) measures the variation of  $f$  in a neighborhood of size proportional to  $a$  centered on point  $b$ . In order, to reconstruct  $f$  from its wavelet transform (4.18), one needs a reconstruction formula, known as the resolution of the identity ([60, 62]).

$$f(t) = \frac{1}{C_\psi} \int_0^\infty \int_{-\infty}^\infty W_f(a,b) \psi_{a,b}(t) \frac{da db}{a^2}. \quad (4.19)$$

From the above equation we can see why the condition given by Eq. 4.15 should be imposed. One fundamental property that we require in order to analyze singular behavior is that  $\psi(t)$  has enough vanishing moments as argued in the works of [57] and [61]. A wavelet is said to have  $n$  vanishing moments if and only if it satisfies

$$\int_{-\infty}^{\infty} t^k \psi(t) dx = 0, \text{ for } k = 0, 1, \dots, n-1 \quad (4.20)$$

and

$$\int_{-\infty}^{\infty} t^k \psi(t) dt \neq 0, \text{ for } k \geq n. \quad (4.21)$$

This means that a wavelet with  $n$  vanishing moments is orthogonal to polynomials up to order  $n - 1$ . In fact, the admissibility condition given by Eq. (4.15) requires at least one vanishing moment. So the wavelet transform of  $f(t)$  with a wavelet  $\psi(t)$  with  $n$  vanishing moments is nothing but a “smoothed version” of the  $n$ -th derivative of  $f(t)$  on various scales. In fact, when someone is interested to measure the local regularity of a signal this concept is crucial (see for instance [60, 62]).

#### 4.5.2 Wavelet singularity analysis

The local regularity of a function  $f$  at a point  $t_0$  is often measured by its Hölder exponent. The Hölder exponent  $\alpha$  measures the strength of a singularity at a particular point  $t_0$ , where  $t_0$  belongs to the domain of  $f$ , see the Appendix. It is important to point out that if the singular part of a function  $f$  in the neighborhood of  $t_0$  is of the type  $|t - t_0|^\alpha$ , then it corresponds to a *cusp* and in this case the singular behavior is fully characterized by its Hölder exponent. However, there exists functions that involve oscillating singularities which have to be described by an additional quantity: an oscillating exponent ([58, 59]). In such a case, the oscillation has to be analyzed carefully. Such functions can not be fully characterized only by the Hölder exponent. In this work, we will only consider functions whose singularities are not oscillating.

One classical tool to measure the regularity of a function  $f(t)$  is to look at the asymptotic decay of its Fourier transform  $\hat{f}(\omega)$  at infinity. However, the Fourier transform is not well adapted to measure the local regularity of functions, because it is global and provides a description of the overall regularity of functions ([61, 62]). Consequently, we need another way to characterize local signal regularity.

In the works [57, 60, 61, 62] it is shown that the WT provides a way of doing a precise analysis of the regularity properties of functions. This is made possible by the scale parameter. Due to its ability to focus on singularities in the signals, the WT is sometimes referred to as ‘mathematical microscope’ ([57, 60, 61, 62]), where the wavelet used determines the optics of the microscope and its magnification is given by the scale factor  $a$ .

The WT modulus maxima (WTMM) decomposition introduced by [61] provides a local analysis of the singular behavior of signals. In the works of Mallat [61, 62] it has been shown that for cusp singularities the location of the singularity can be detected and the related exponent can be recovered from the scaling of the WT along the so-called *maxima line* (WTMML for short), which is convergent towards the singularity. This is a line where the WT reaches local maximum with respect to the position coordinate. Connecting such local maxima within the continuous WT ‘landscape’ gives rise to the entire tree of maxima lines. Restricting oneself to the collection of such maxima lines provides a particularly useful representation of the entire WT. It incorporates the main characteristics of the WT: the ability to reveal the *hierarchy* of (singular) features, including the scaling behavior.



An other key concept, in addition to vanishing moments, used to characterize the regularity of a function in terms of WTMM is given next. Suppose that  $\psi$  has compact support  $[-C, C]$ . The *cone of influence* of  $\psi$  at point  $t_0$  is the set of points  $(a, b)$  in the scale-space plane or domain, such that  $t_0$  is in the support of  $\psi_{a,b}(t)$ . We will denote the scale-space plane or domain of the WT as the  $(a, b)$ -plane or the  $(a, b)$ -domain. Since the support of  $\psi((t-b)/a)$  is  $[b-Ca, b+Ca]$ , the point  $(a, b)$  belongs to the cone of influence of  $t_0$  if

$$|b - t_0| \leq Ca. \quad (4.22)$$

The function  $f(t)$  has a Hölder exponent  $\alpha \in (k, k+1)$  at  $t_0$ , if and only if there exists a constant  $A > 0$  such that at each modulus maxima  $(a, b)$  in the cone defined by Eq. (4.22) one has

$$|W_f(a, b)| \leq Aa^{\alpha+1/2}, \quad a \rightarrow 0, \quad (4.23)$$

(see [61, 62]). Here it is assumed that the wavelet has at least  $n > \alpha$  vanishing moments. If  $f(t)$  is regular at  $t_0$  or, if the number of vanishing moments is too small, i.e.,  $n < \alpha$ , one obtains for  $a \rightarrow 0$  a scaling behavior of the type

$$|W_f(a, b)| \leq Aa^{n+1/2}. \quad (4.24)$$

The scaling behavior of the WTMM is given in Eq. (4.23) and can be rewritten as follows

$$\log |W_f(a, b)| \leq \log A + \left( \alpha + \frac{1}{2} \right) \log a. \quad (4.25)$$

The global Hölder regularity at  $t_0$  is thus the maximum slope  $-\frac{1}{2}$  of  $\log |W_f(a, b)|$  as a function of  $\log a$  along the maxima line converging to  $t_0$ .

## 4.6 Results and discussion

In this section, we present the results we obtained using the singularity detection procedure described in the previous section. The signal to be analyzed,  $f(t) = y$ , represents the evolution in time of the total biomass concentration for the fermentation processes described in Sections 4.3 and 4.4 that includes different cases as specified therein. In all the wavelet-related calculations we employed as mother wavelets the first and second derivative of the Gaussian function, having one and two vanishing moments, respectively. The final goal is always to calculate the Hölder exponent of the singularities for such processes because it is a direct measure of the irregularity of a signal (function) at the singular point  $t_0$ , in the sense that higher values of it correspond to more regular functions than the lower values.

Figure 1 *a,b,c* shows the performance of the wavelet singularity analysis as applied to Case I (same initial conditions but different kinetic rates). We obtain a Hölder exponent of quite high value.

The following figure shows the performance of the scheme applied to Case II (same growth rates but different initial conditions). In this case, the Hölder exponent of the mixed growth singularity is lower than in Case I.

Similarly to the previous cases, Fig. (3) presents the graphical results for Case III (different initial conditions and different growth rates). Although, the singularity looks very mild in the time evolution of the total biomass concentration the wavelet analysis is able to detect it with high precision.

Finally, Case IV (same initial conditions, same growth rates but with different values of their maximal growth rates) is graphically analyzed in Fig. (4). For this case we obtained the lowest Hölder exponent.

Although the latter two cases seem to correspond to almost overlapping of the WTMMML pointing to bifurcation phenomena we are still not at the threshold of a completely different behavior of the log plots generated by bifurcations. This could be explained by the fact that the strength of the first singularity is bigger with respect to the second one.

#### 4.6.1 Wavelet analysis for the MCMS case

The MCMS case is the most interesting case that we discuss here because we will show that it is possible to infer in a very accurate manner by means of WT the moment in which the microorganisms switch their carbon source. In order to understand the detailed dynamics of this combined growth, we first apply separately the WT approach to the two biomass signals  $y = c_1$  (Fig. 5) and  $y = c_2$  (Fig. 6) and then to the total signal  $y = c_1 + c_2$  (Fig. 7).

It is worth noting that the Hölder exponent is bigger than one, a quite interesting feature which means that the singularity lies in the second derivative of the biomass signal. This result gives further opportunities to characterize the nature of the singularity because it suggests that the type of growth can be inferred from the order of the derivative in which the singularity occurs directly given by the value of the Hölder exponent. The latter fact is a great simplification with respect to the analytical search of the singularities which implies obtaining the analytical solution of the given dynamical growth model. Moreover, even in such fortunate cases, the analytical solutions could be subject to fixed parameter values of the model. On the other hand, the WT numerical approach allows the singularity analysis even in the case of time-varying parameters.

#### 4.6.2 Wavelet analysis for the noisy data case

It is well known that in some cases the amplitude of the Gaussian noise affecting the on-line signals can be an important annoying factor. Therefore, a good analysis should be robust in such cases. Thus, we provide here the WT analysis for the MCMS biomass signals in the presence

of white noise, that is, in the next figures (8-10) we consider signals of the form  $y_i = c_i + \varepsilon(t)$  and their sum, where  $\varepsilon(t)$  stands for the functional form of the noise.

We notice from the corresponding plots that the noisy data do not allow to obtain global Hölder exponents in a straightforward manner, which is a result already reported in the wavelet literature ([62] and [63]). On the other hand, the singularity detection is robust with respect to the noise for reasonable levels of its amplitude. In addition, Figure 9 gives us the hint that when the cones of influence produced by the Gaussian noise enter the scales of the singularity the cone of the latter becomes undistinguishable from those of the noise. This remark could be used as a sort of resolution criterium of the WT method in the presence of noise. Therefore, one can determine a critical amplitude of the noise for which the WT approach loses its applicability.

## 4.7 Concluding remarks

We showed here explicitly how the wavelet singularity analysis can be applied to infer mixed growth behavior of fermentation processes using only total biomass data. We prove that the wavelet analysis is very accurate for all the cases we considered. A very interesting feature of our research is that the Hölder exponent is sensitive to the type of the mixed-growth phenomenon, more specifically depends on the parameters of the growth processes and on their initial conditions. The MCMS case points to the remarkable technological possibility of detecting the change of the substrate uptake since the singularity appears in the second derivatives of the biomass signal. This can lead one to think of the possibility to infer substrate contaminations based only in the analysis of the biomass data. In addition, our results for the noisy data clearly hint to the fact that the wavelet singularity analysis maintains its attractive features even in these more difficult but realistic case. We hope that in future works we could find out the mathematical relationships implied by this possible correlation. It might allow the usage of the Hölder exponent as an identification criterium of the more specific nature of mixed-growth processes.

## Appendix

A function  $f : \mathbb{R} \rightarrow \mathbb{R}$  is said to be Hölder continuous of exponent  $\alpha$  ( $0 < \alpha < 1$ ) if, for each bounded interval  $(c, d) \subset \mathbb{R}$ , we can find a positive constant  $K$  such that

$$|f(t) - f(t_0)| \leq K|t - t_0|^\alpha \quad (4.26)$$

for all  $t, t_0 \in (c, d)$ .

The space of Hölder continuous functions is denoted  $C^\alpha$ . A function is said to be  $C^{n+\alpha}$  if it is in  $C^n$  and its  $n$ th derivative is Hölder continuous with exponent  $\alpha$ . Thus, if we consider the Hölder exponent  $n < \alpha < n + 1$ , with  $n \in \mathbb{N}$ , the function can be differentiated  $n$  times, but the  $(n + 1)$ th derivative does not exist. Therefore, a function with a Hölder exponent  $n < \alpha < n + 1$  is said to be singular in the  $n$ th derivative. Keeping this in mind, let us give the following definition of the Hölder regularity of a function [60, 61, 62].

- Let  $n \in \mathbb{N}$  and  $n \leq \alpha < n + 1$ . A function  $f(t)$  has a *local* Hölder exponent  $\alpha$  at  $t_0$  if and only if there exist a constant  $K > 0$ , and a polynomial  $P_n(t)$  of order  $n$ , such that

$$\forall t \in \mathbb{R}, \quad |f(t) - P_n(t - t_0)| \leq K|t - t_0|^\alpha \quad (4.27)$$

- The function  $f(t)$  has a *global* Hölder exponent  $\alpha$  on the interval  $(c, d)$  if and only if there is a constant  $K$  and a polynomial of order  $n$ ,  $P_n(t)$ , such that equation (4.27) is satisfied for all  $t \in (c, d)$ .
- The Hölder *regularity* of  $f(t)$  at  $t_0$  is the supremum of the  $\alpha$  such that  $f(t)$  is Hölder  $\alpha$  at  $t_0$ .
- The  $n$ th derivative of a function  $f(t)$  is *singular* at  $t_0$  if  $f(t)$  has a local Hölder exponent  $\alpha$  at  $t_0$  with  $n < \alpha < n + 1$ .

A function  $f(t)$  that is continuously differentiable at a given point has a Hölder exponent not less than 1 at this point. If  $\alpha \in (n, n + 1)$  in (4.27) then  $f(t)$  is  $n$  times but not  $(n + 1)$  times differentiable at the point  $t_0$ , and the polynomial  $P_n(t)$  corresponds to the first  $(n + 1)$  terms of the Taylor series of  $f(t)$  around  $t = t_0$ . For example, if  $n = 0$ , we have  $P_0(t - t_0) = f(t_0)$ .

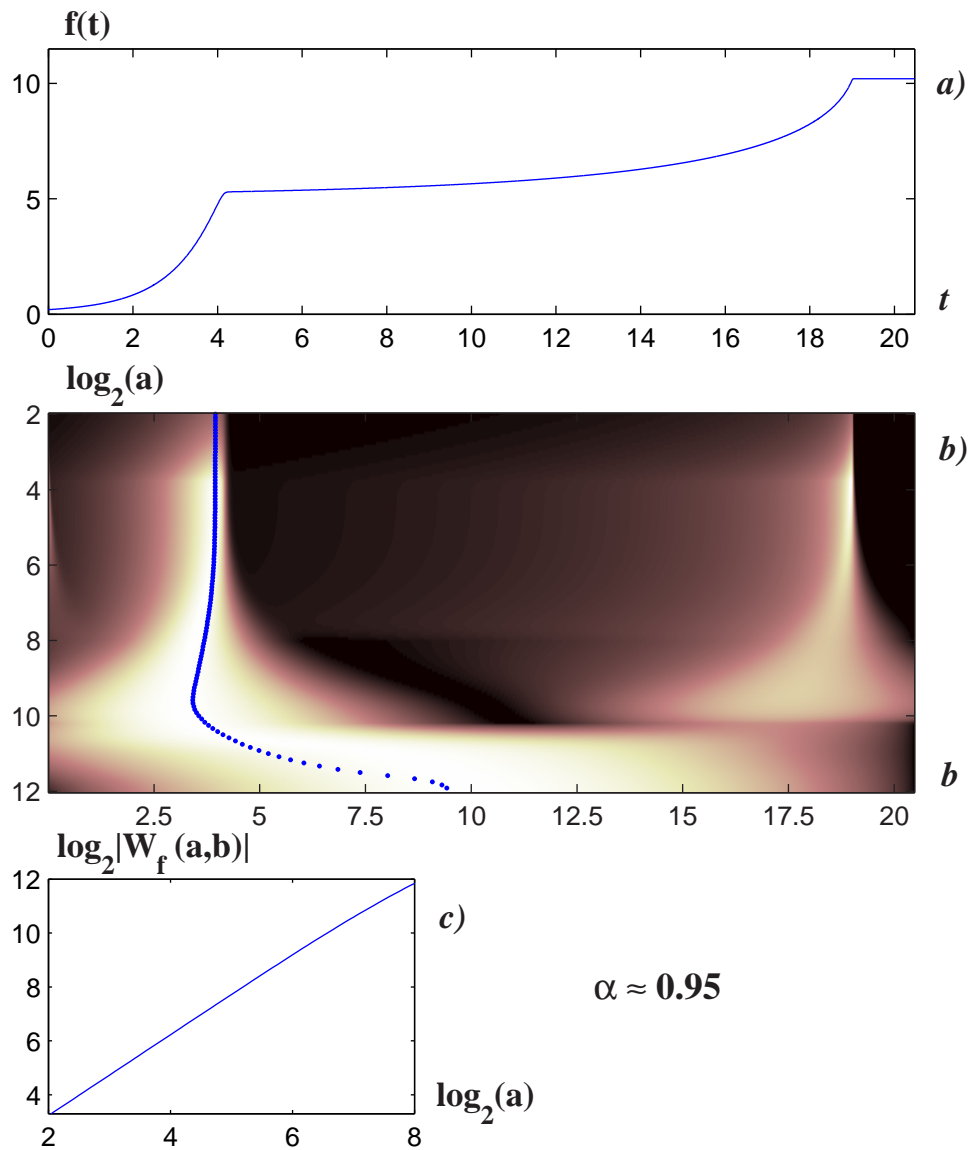


Figure 4.1: *a)* The time evolution of the total biomass concentration signal for Case I. *b)* The wavelet cones of influence corresponding to this case showing a very accurate identification of the two singularity points presented in the signal, of which the first one allows to infer the presence of the mixed growth feature of the fermentation process whereas the second one is associated with the end of the fermentation batch cycle. *c)* From the slope in the double logarithmic plot, the Hölder coefficient of the mixed growth singularity is calculated as  $\alpha = 0.95$ .

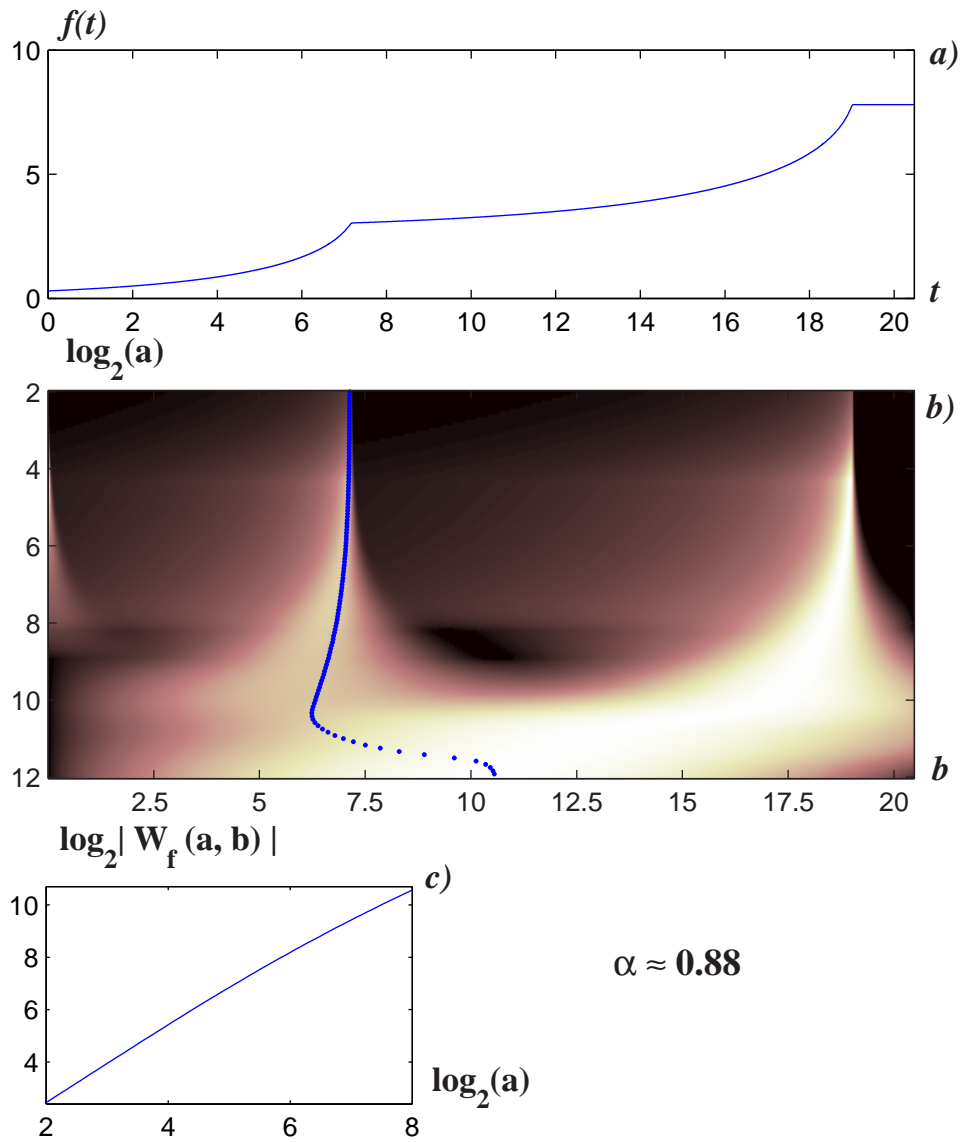


Figure 4.2: *a)* The time evolution of the total biomass concentration signal for Case II. *b)* The wavelet cones of influence corresponding to this case again showing the accurate identification of the two singularity points, of the same type, respectively, as in Fig. 1. *c)* The Hölder coefficient of the mixed growth singularity is now  $\alpha = 0.88$ .

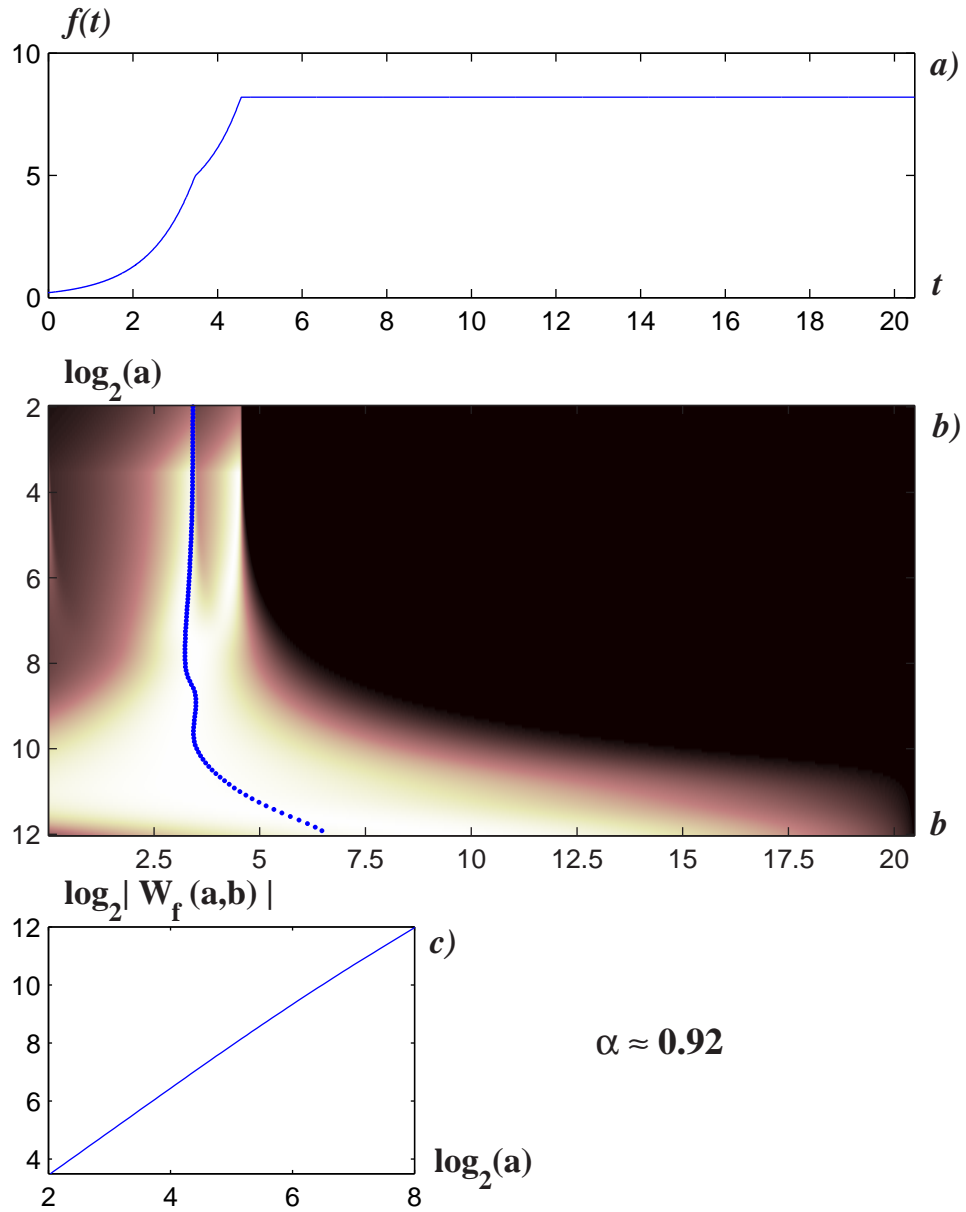


Figure 4.3: Same caption comments as in the previous figures but for Case III. The Hölder coefficient of the mixed growth singularity is now  $\alpha = 0.92$ .

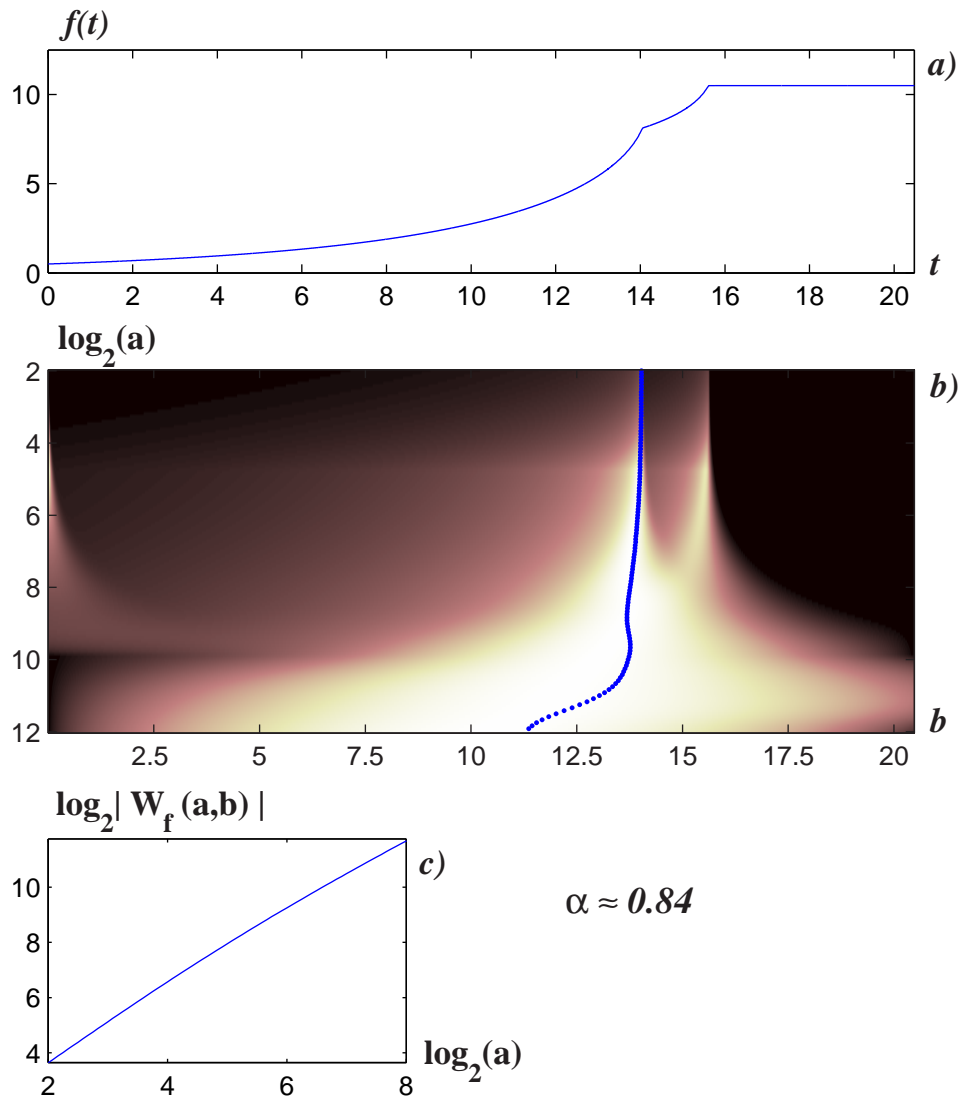


Figure 4.4: Same caption comments as in the previous figures but for Case IV. The value of the Hölder coefficient for the mixed growth singularity is  $\alpha = 0.84$ .



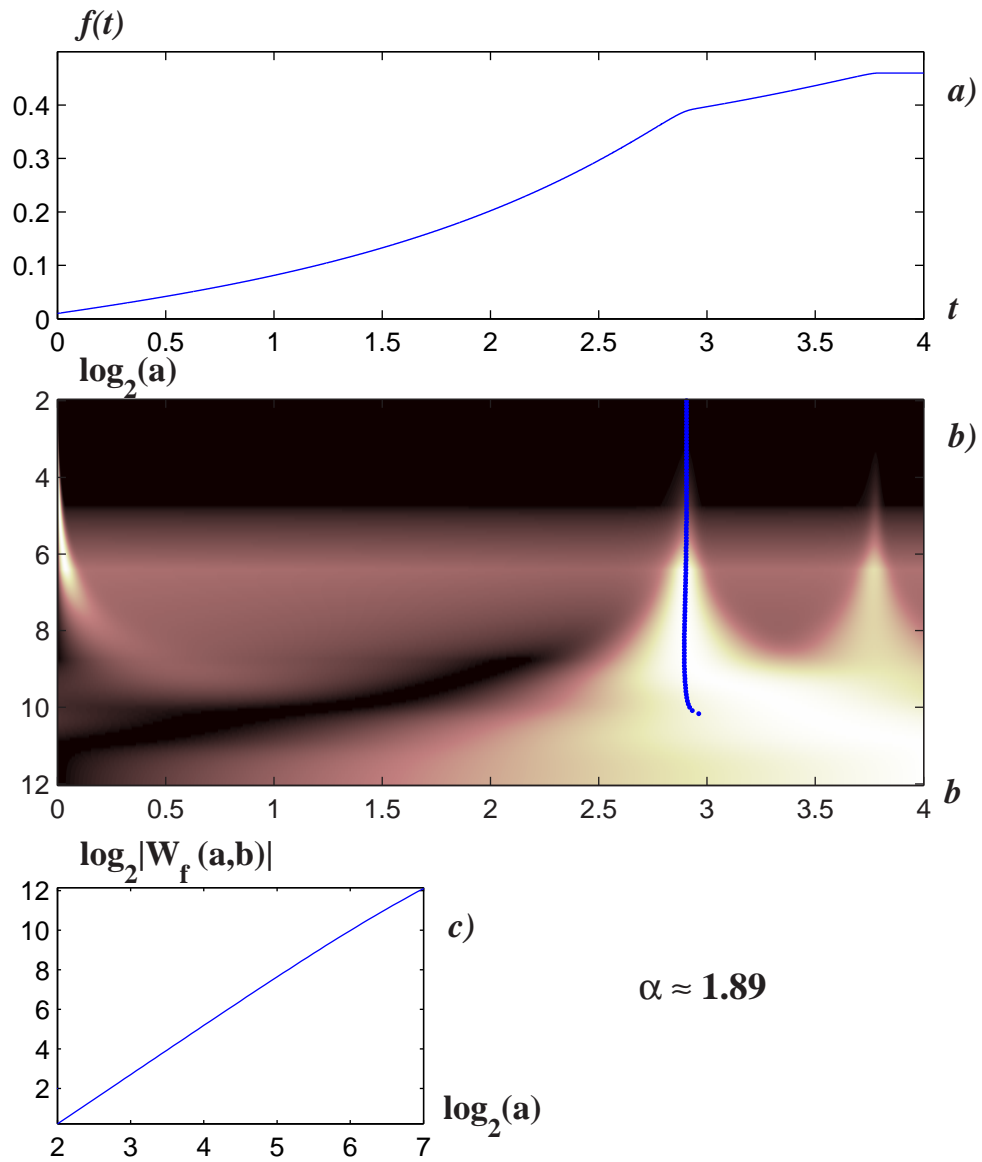


Figure 4.5: Same caption comments as in the previous figures but for the MCMS biomass signal  $y = c_1$ . The Hölder coefficient of the mixed growth singularity is now  $\alpha = 1.89$ .

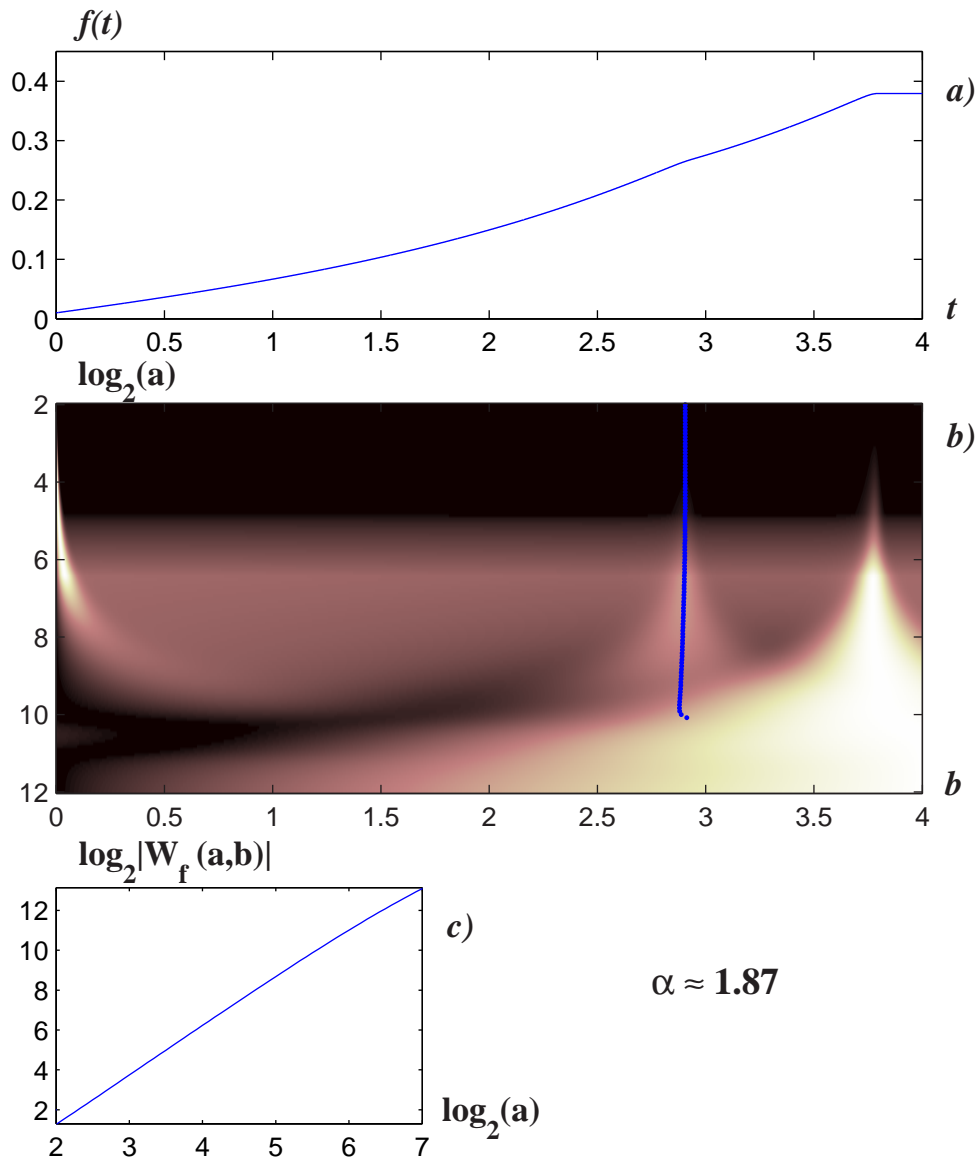


Figure 4.6: Same caption as in the previous figures but for the MCMS biomass signal  $y = c_2$ . The Hölder coefficient of the mixed growth singularity is now  $\alpha = 1.87$ .

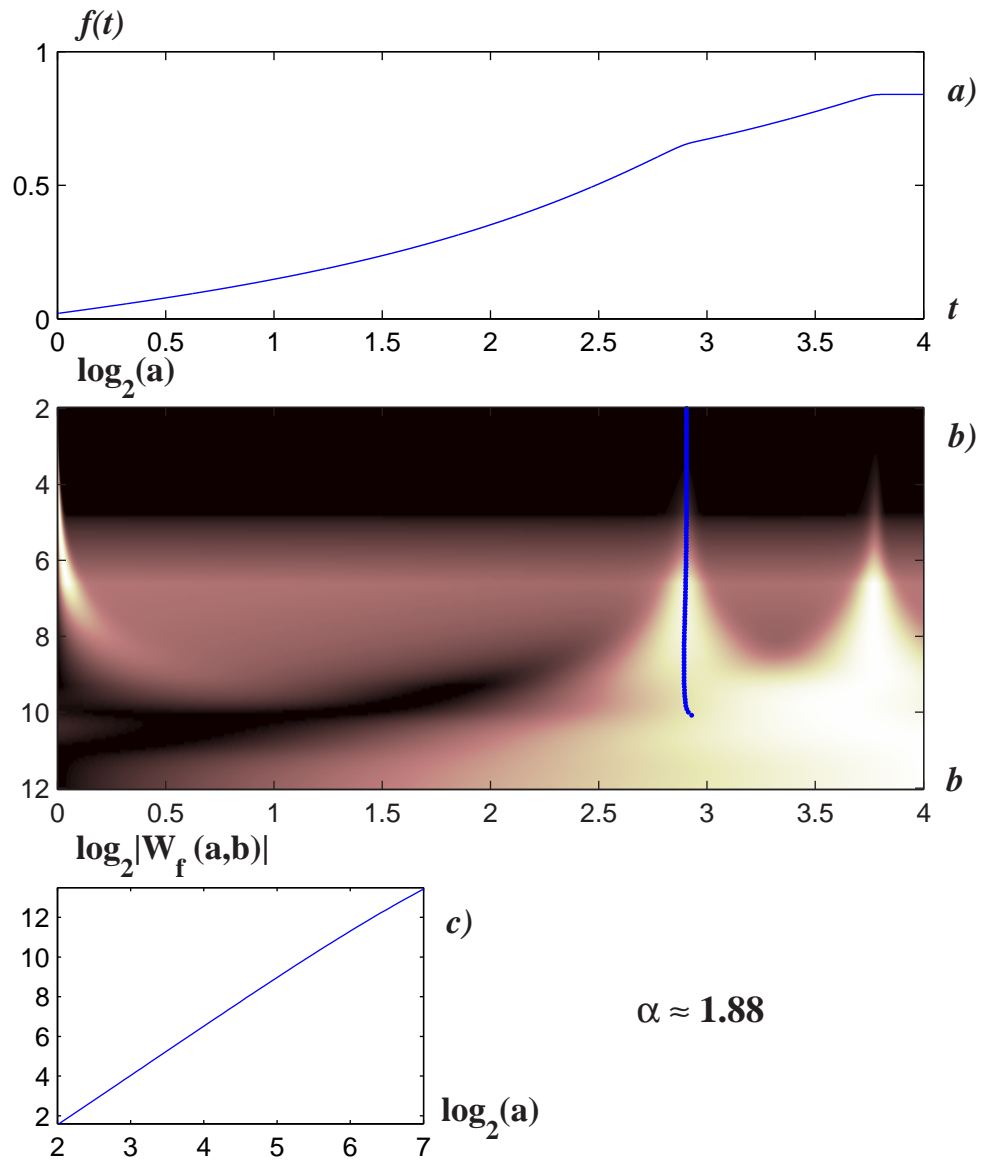


Figure 4.7: Same caption comments as in the previous figures but for the total MCMS signal. The Hölder coefficient of the mixed growth singularity is now  $\alpha = 1.88$ .

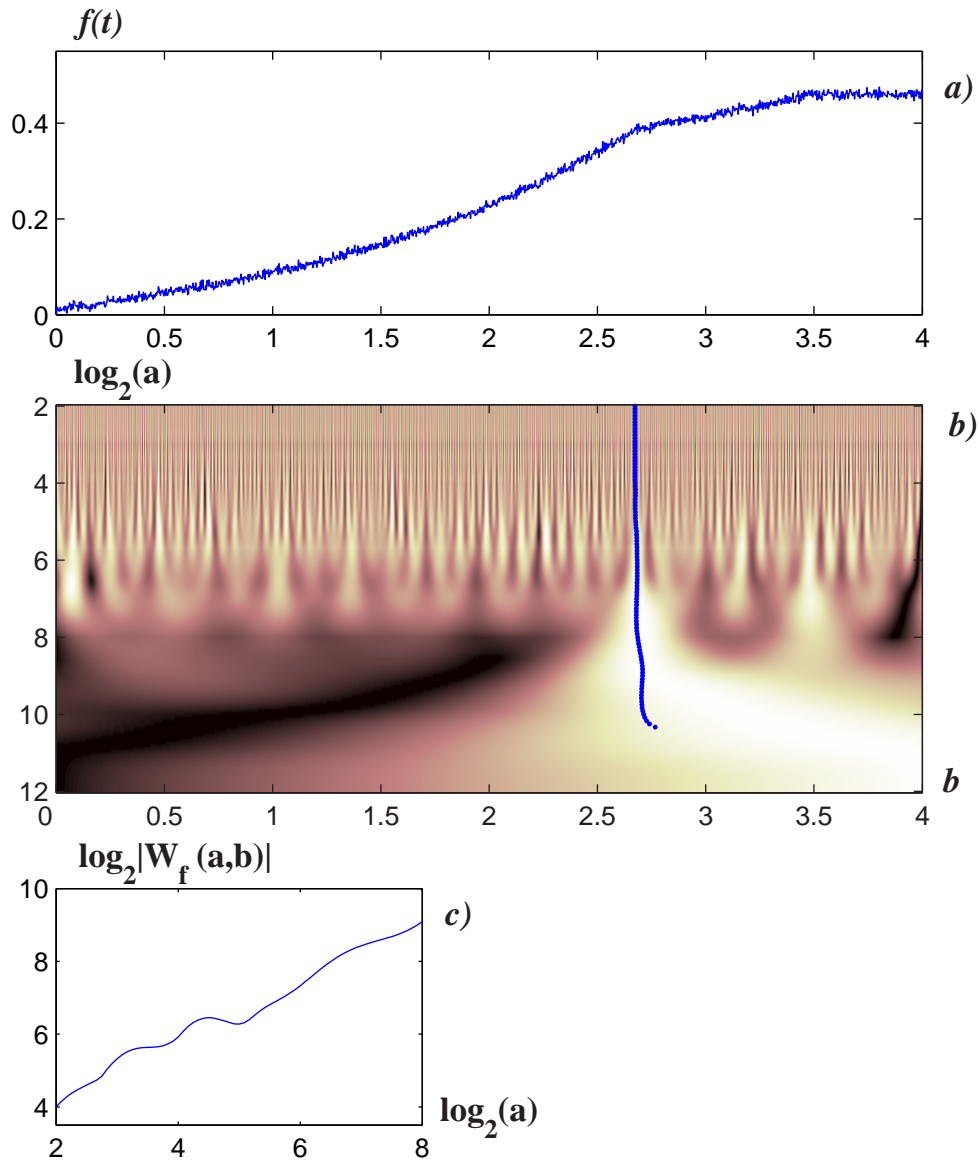


Figure 4.8: MCMS data corresponding to  $c_1$  with a small amplitude Gaussian noise added. From the bottom plot c) one can see that because the curve is not a straight line one cannot get a global Hölder coefficient from its slope.

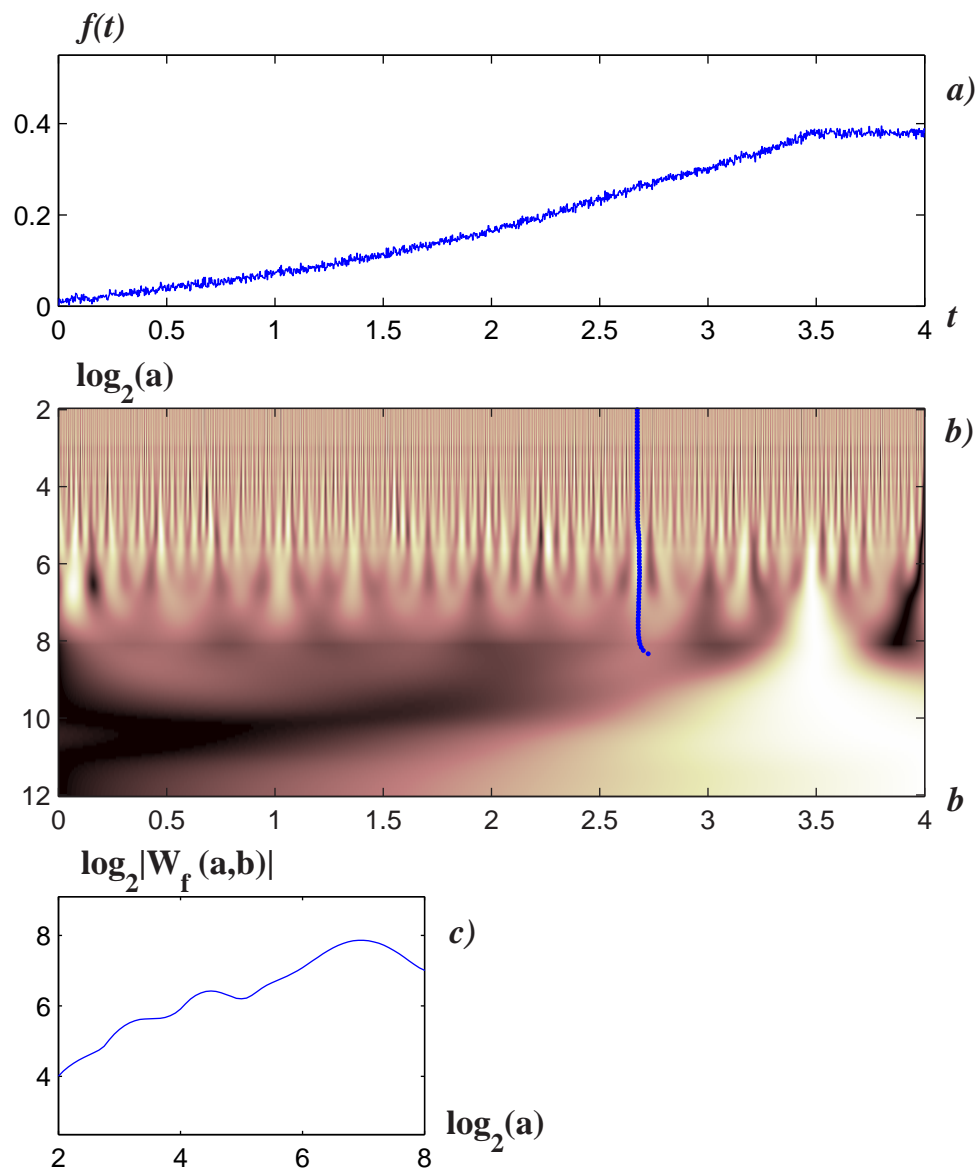


Figure 4.9: MCMS data corresponding to  $c_2$  with a small amplitude Gaussian noise added. The Hölder coefficient is not a useful concept in this case.

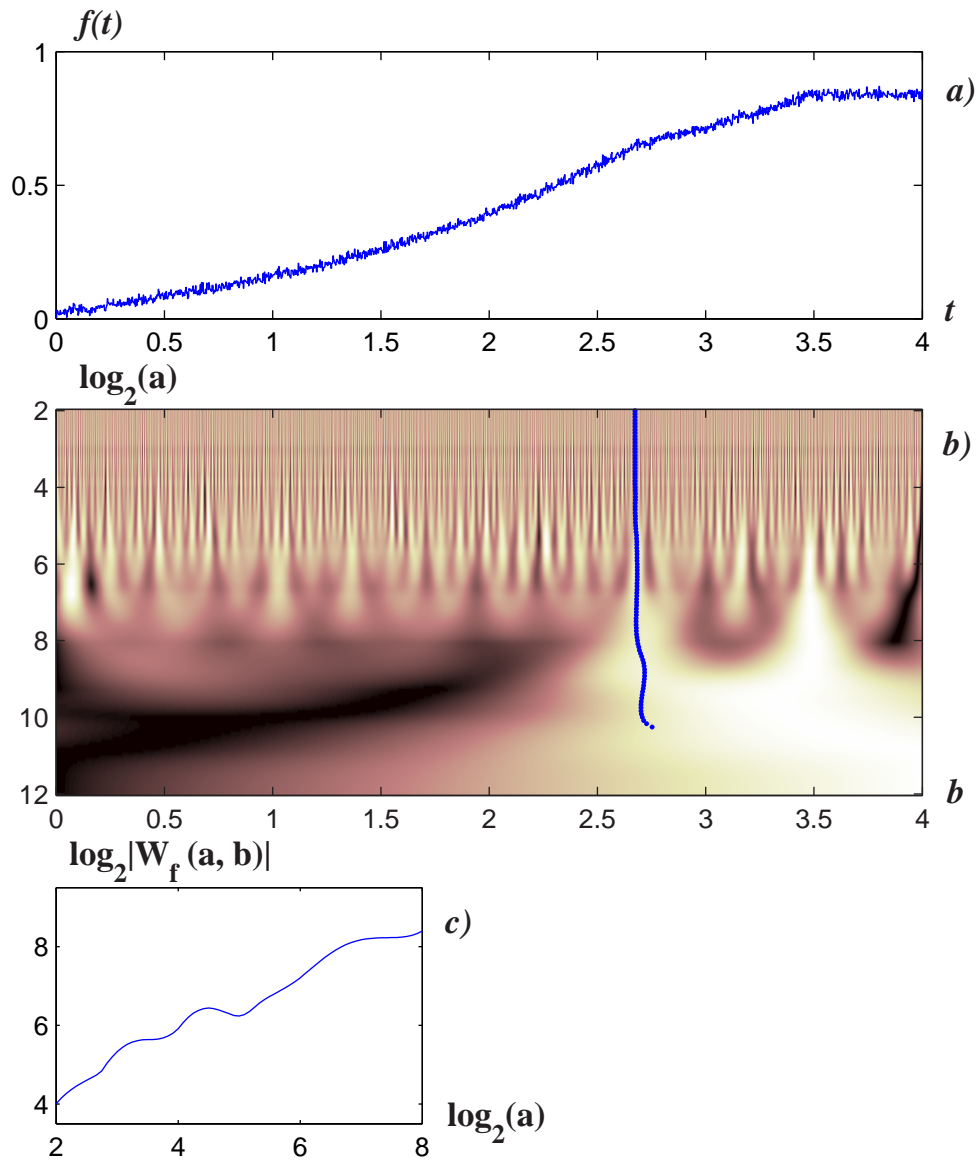


Figure 4.10: MCMS data corresponding to the sum  $c_1 + c_2$  with the same Gaussian noise added. The concept of global Hölder coefficient is again not useful.

---

---

## Gene Regulation Networks

---

### 5.1 Nonlinear Software Sensor for Monitoring Genetic Regulation Processes with Noise and Modeling Errors

This work was performed in collaboration with the Ph.D. student Luis Adolfo Torres, Professor Dr. H. C. Rosu, Professor Dr. Gerardo Argüello, and Professor Dr. Julio Collado-Vides. [64].

Nonlinear control techniques by means of a software sensor that are commonly used in chemical engineering could be also applied to genetic regulation processes. We provide here a realistic formulation of this procedure by introducing an additive white Gaussian noise, which is usually found in experimental data. Besides, we include model errors, meaning that we assume we do not know the nonlinear regulation function of the process. In order to illustrate this procedure, we employ the Goodwin dynamics of the concentrations [B.C. Goodwin, *Temporal Oscillations in Cells*, (Academic Press, New York, 1963)] in the simple form recently applied to single gene systems and some operon cases [H. De Jong, *J. Comp. Biol.* **9**, 67 (2002)], which involves the dynamics of the mRNA, given protein, and metabolite concentrations. Further, we present results for a three gene case in co-regulated sets of transcription units as they occur in prokaryotes. However, instead of considering their full dynamics, we use only the data of the metabolites and a designed software sensor. We also show, more generally, that it is possible to rebuild the complete set of nonmeasured concentrations despite the uncertainties in the regulation function or, even more, in the case of not knowing the mRNA dynamics. In addition, the rebuilding of concentrations is not affected by the perturbation due to the additive white Gaussian noise and also we managed to filter the noisy output of the biological system.

## 5.2 Introduction

Gene expression is a complex dynamic process with intricate regulation networks all along its stages leading to the synthesis of proteins [65]. Currently, the most studied aspect is that of regulation of initiation of transcription at the DNA level. Nevertheless, the expression of a gene product may be regulated at several levels, from transcription to RNA elongation and processing, RNA translation and even as post-translational modification of protein activity. Control engineering is a key discipline with tremendous potential to simulate and manipulate the processes of gene expression. In general, the control terminology and its mathematical methods are poorly known to the majority of biologists. Many times the control ideas are simply reduced to the homeostasis concept. However, the recent launching of the IEE journal *Systems Biology* [66] points to many promising developments from the standpoint of systems analysis and control theory in biological sciences. Papers like that of Yi et al [67], in which the Barkai and Leibler robustness model [68] of perfect adaptation in bacterial chemotaxis is shown to have the property of a simple linear integral feedback control, could be considered as pioneering work in the field.

We mention here two important issues. The first one is that the basic concept of state of a system or process could have many different empirical meanings in biology. For the particular case of gene expression, the meaning of a state is essentially that of a concentration. The typical problem in control engineering that appears to be tremendously useful in biology is the reconstruction of some specific regulated states under conditions of limited information. Moreover, equally interesting is the issue of noise filtering. It is quite well known that gene expression is a phenomenon with two sources of noise: one due to the inherent stochastic nature of the process itself and the other originating in the perturbation of the natural signal due to the measuring device. In the mathematical approach, the latter class of noise is considered as an additive contamination of the real signal and this is also our choice here. Both issues will form the subject of this investigation.

Taking into account the fact that rarely one can have a sensor on every state variable, and some form of reconstruction from the available measured output data is needed, a software can be constructed using the mathematical model of the process to obtain an estimate  $\hat{X}$  of the true state  $X$ . This estimate can then be used as a substitute for the unknown state  $X$ . Ever since the original work by Luenberger [69], the use of state observers has proven useful in process monitoring and for many other tasks. We will call herein as observer, in the sense of control theory, an algorithm capable of giving a reasonable estimation of the unmeasured variables of a process. For this reason, it is widely used in control, estimation, and other engineering applications.

Since almost all observer designs are heavily based on mathematical models, the main drawback



is precisely the dependence of the accuracy of such models to describe the naturally occurring processes. Details such as model uncertainties and noise could affect the performance of the observers. Taking into account these details is always an important matter and should be treated carefully. Thus, we will pay special attention in this research to estimating unknown states of the gene expression process under the worst possible case, which corresponds to noisy data, modeling errors, and unknown initial conditions. These issues are of considerable interest and our approach is a novel contribution to this important biological research area. Various aspects of noisy gene regulation processes have been dealt with recently from both computational and experimental points of view in a number of interesting papers [70]. We point out that since we add the noise  $\delta$  to the output of the dynamic system in the form  $y = CX + \delta$  (see Eqs.  $\Gamma$  in Section IV) it seems that its origin is mainly extrinsic to the regulation process, even though it could be considered as a type of intrinsic noise with respect to the way the experiment is performed. On the other hand, when writing the equation in the form  $y = C(X + I\Delta)$ , where  $\Delta$  is a vector of noisy signals, one can see that the observer could estimate states that are intrinsically noisy even though the processes are still deterministic.

### 5.3 Brief on the biological context

Similar to many big cities, with heavy traffic, biological cells host complicated traffic of biochemical signals at all levels. Like cars on a busy highway, millions of molecules get involved in the bulk of the cell in many life processes controlled by genes. At the nanometer level, clusters of molecules in the form of proteins drive the dynamics of the cellular network that schematically can be divided into four regulated parts: the DNA or genes, the transcribed RNAs, the set of interacting proteins and the metabolites. Genes can only affect other genes through specific proteins, as well as through some metabolic pathways that are regulated by proteins themselves. They act to catalyze the information stored in DNA, all the way from the fundamental processes of transcription and translation to the final quantities of produced proteins.

Considering the enormous complexity of multicellular organisms generated by their large genomes, one can nevertheless still associate at least one regulatory element to any component gene. Each regulatory system is then composed of two elements at the DNA level, the gene that encodes a transcriptional regulator, and the target in the DNA where this regulator binds to, and exerts its activator or repressor function in transcription. These loops of interactions represent a fundamental piece to understand the functioning of complex regulatory transcriptional and translational networks [71, 72]. For the purpose of modelling, it is essential to generate simple models that help to understand elementary dynamical components of these complex regulatory networks as molecular tools that participate in an important way in the machinery of cellular decisions, that is to say, in the behaviour and genetic program of cells.

Many entities in cellular networks can be identified as the basic units of regulation, mainly distinguished by their unique roles with respect to interaction with other units. These basic units are: the genes, with codifying content, also described as structural genes; the regulatory elements that in the old literature were called regulatory genes, which are smaller fragments of DNA sequences (of the order of 5 to 20 nucleotides) called operator sites where regulatory proteins as well as the RNA polymerase bind to; the messenger RNAs or mRNAs which are the products of transcription and form the template for the subsequent production of proteins as encoded by the corresponding gene; the forms of each protein and protein complexes, as well as, all metabolites present in the cell, either as products of enzymatic reactions or internalized by transport systems. These units have associated values that either represent concentrations or levels of activation. These values depend on both the values of the units that affect them due to the aforementioned mechanisms and on some parameters that govern each special form of interaction.

This gives rise to genetic regulatory systems structured by networks of regulatory interactions between DNA, RNA, proteins, and small molecules. The simplest regulatory network is made of only one gene that is transcribed into mRNA, this mRNA is then translated into proteins, which can be activated or inhibited as a result of their interaction with other proteins or with specific metabolites. Transcriptional regulators are two-head structures, one being the domain of DNA interaction, and the other one is the so-called allosteric domain that interacts with specific metabolites. Taking together these properties of the molecular machinery, one can envision that a gene encodes a protein which can regulate its own activity, either positively or negatively, depending on its effect in enhancing or preventing the RNA polymerase transcriptional activity on its own gene by means of binding to an operator sites upstream of its own encoding gene. Upstream here meaning before the beginning of the gene where transcription initiates. A mathematical model of such a biological inhibitory loop has been discussed since a long time ago by Goodwin and recurrently occurred in the literature, most recently being reformulated by De Jong [73]. Although this case could look unrealistic, there are simple organisms, such as bacteria, where one regulatory loop may prove essential as recently discussed in detail by Ozbudak et al [74]. However, already at the level of two genes the situation gets really complicated, mostly because of the possible formation of heterodimers between the repressors and other proteins around. These heterodimers are able to bind at the regulatory sites of the gene and therefore can affect it and lead to modifications of the regulatory process.

Recent development of experimental techniques, like cDNA microarrays and oligonucleotide chips, have allowed rapid measurements of the spatiotemporal expression levels of genes [75, 76, 77]. In addition, formal methods for the modeling and simulation of gene regulation processes are currently being developed in parallel to these experimental tools. As most genetic regulatory systems of interest involve many genes connected through interlocking positive and negative feedback loops, an intuitive understanding of their dynamics is hard to obtain. The advantage of the formal methods is that the structure of regulatory systems can be described

unambiguously, while predictions of their behavior can be made in a systematic way.

To make the description very concrete, it is interesting to look at well-defined, i.e., quite simple mathematical models that we present in the next section that refers to single gene cases and single gene clusters (operons). The nonlinear software sensor for such cases is discussed in Section IV. A three-gene case is treated as an extension to regulatory gene networks and shows that the method of forward engineering still works for reasonably simple gene networks. The conclusion section comes at the end of the chapter.

## 5.4 Mathematical Model for Gene Regulation

In this section, we use the very first kinetic model of a genetic regulation process developed by Goodwin in 1963 [79], generalized by Tyson in 1978 [80] and most recently explained by De Jong [73]. The model in its most general form is given by the following set of equations:

$$\dot{X}_1 = K_{1n}r(X_n) - \gamma_1 X_1, \quad (5.1)$$

$$\dot{X}_i = K_{i,i-1}X_{i-1} - \gamma_i X_i, \quad 1 < i \leq n. \quad (5.2)$$

The parameters  $K_{1n}, K_{21}, \dots, K_{n,n-1}$  are all strictly positive and represent production constants, whereas  $\gamma_1, \dots, \gamma_n$  are strictly positive degradation constants. These rate equations express a balance between the number of molecules appearing and disappearing per unit time. In the case of  $X_1$ , the first term is the production term involving a nonlinear nondissipative regulation function. We take this as an unknown function. On the other hand, the concentration  $X_i$ ,  $1 < i \leq n$ , increases linearly with  $X_{i-1}$ . As well known, in order to express the fact that the metabolic product is a co-repressor of the gene, the regulation function should be a decreasing function for which most of the authors use the Hill sigmoid, the Heaviside and the logoid curves. The decrease of the concentrations through degradation, diffusion and growth dilution is taken proportional to the concentrations themselves. For further details of this regulation model we recommend the reader the review of De Jong [73].

It is to be mentioned here that bacteria have a simple mechanism for coordinating the regulation of genes that encode products involved in a set of related processes: these genes are clustered on the chromosome and are transcribed together. Most prokaryotic mRNAs are polycistronic (multiple genes on a single transcript) and the single promoter that initiates transcription of clusters is the site of regulation for expression of all genes in the cluster. The gene cluster and promoter, plus additional sequences that function together in regulation, are called operon. Operons that include two to six genes transcribed as a unit are common in nature [81].

The fact that two or more genes are transcribed together on one polycistronic mRNA implies that we have a unique mRNA production constant and consequently we also have one mRNA

degradation constant. In addition, the polycistronic mRNA can be translated into one or several enzymes, resulting in the existence of just one enzyme production and degradation constant, respectively. The same applies for the metabolite produced through the enzyme catalysis. Thus, if the resulting metabolite has repressor activity over the polycistronic mRNA (as in the case of tryptophan [82]), then the model given by Eqs. (5.1,5.2) could also be applied to operons and therefore it has a plausible application to the study of prokaryotic gene regulation.

## 5.5 Nonlinear Software Sensor

Numerous attempts have been made to develop nonlinear observer design methods. One could mention the industrially popular extended Kalman filter, whose design is based on a local linearization of the system around a reference trajectory, restricting the validity of the approach to a small region in the state space [78, 83]. The first systematic approach for the development of a theory of nonlinear observers was proposed some time ago by Krener and Isidori [84]. In further research, nonlinear transformations of the coordinates have also been employed to put the considered nonlinear system in a suitable “observer canonical form”, in which the observer design problem may be easily solved [85, 86, 87]. Nevertheless, it is well known that classical proportional observers tend to amplify the noise of on-line measurements, which can lead to the degradation of the observer performance. In order to avoid this drawback, this observer algorithm is based on the works of Aguilar et al. [88, 89], because the proposed integral observer provides robustness against noisy measurement and uncertainties. We show that this new structure retains all the characteristics of the popular (the traditional high gain) state observers of the classical literature and furthermore provides additional robustness and noise filtering and thus can result in a significant improvement of the monitoring performances of the genetic regulation process.

In this section, we present the design of a nonlinear software sensor in which one  $X_j$ , for  $j \in (1, \dots, n)$ , is the naturally measured state (the most easy to measure). Therefore, it seems logical to take  $X_j$  as the output of the system

$$y = h(X) = X_j . \quad (5.3)$$

Now, considering the constant  $K_{1n}$  and the function  $r(X_n)$  as unknown, we group them together in a function  $\mathfrak{F}(X)$ . In addition, we consider that the output function  $h(X)$  is contaminated with a Gaussian noise. In such a case, the model given by the aforementioned Eqs. (5.1) and (5.2), acquires the form:

$$\Gamma: \begin{cases} \dot{X} = \bar{\mathfrak{F}}(X) + \ell(X) \\ y = CX + \delta \end{cases} \quad (5.4)$$

where  $\bar{\mathfrak{F}}(X)$  is a  $n \times 1$  vector whose first entry is  $\mathfrak{F}(X)$  and all the rest are zero,  $\ell(X)$  is also a  $n \times 1$  vector of the form  $[-\gamma_1 X_1, K_{i,i-1} X_{i-1} - \gamma_i X_i]^T$ ,  $\delta$  is an additive bounded measurement

noise, and  $X \in \mathbb{R}^n$ . The system is assumed to lie in a “physical subset”  $\Sigma \subset \mathbb{R}^n$ .

Then, the task of designing an observer for the system  $\Gamma$  is to estimate the vector of states  $X$ , despite of the unknown part of the nonlinear vector  $\tilde{\mathfrak{S}}(X)$  (which should be also estimated) and considering that  $y$  is measured on-line and that the system is observable. A particular representation of the software sensor that we describe here is provided in Fig. 5.1.

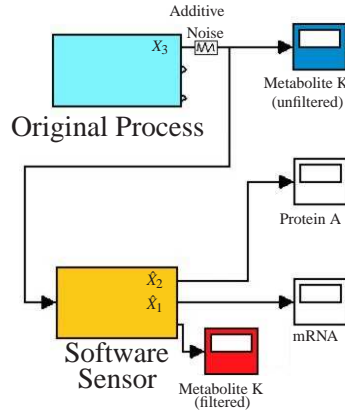


Figure 5.1: Schematic representation of the software sensor, where the output of the system is the input of the software sensor and the outputs of the latter are the rebuilt concentrations.

In order to provide the observer with robust properties against disturbances, Aguilar and collaborators [88] considered only an integral type contribution of the measured error. Moreover, an uncertainty estimator is introduced in the methodology of observation with the purpose of estimating the unknown components of the nonlinear vector  $\tilde{\mathfrak{S}}(X)$ . As a result, the following representation of the system is proposed

$$\Xi : \begin{cases} \dot{X}_0 = CX + \delta \\ \dot{X} = \tilde{\mathfrak{S}} + \ell(X) \\ \dot{\tilde{\mathfrak{S}}} = \Theta(X) \\ y_0 = X_0 \end{cases} \quad (5.5)$$

that is, in the case of the model given by Eqs. (5.1) and (5.2)

$$\dot{X}_0 = X_j + \delta \quad (5.6)$$

$$\dot{X}_1 = X_{n+1} - \gamma_1 X_1 \quad (5.7)$$

$$\dot{X}_i = K_{i,i} X_{i-1} - \gamma_i X_i, \quad 1 < i \leq n, \quad (5.8)$$

$$\dot{X}_{n+1} = \Omega(X) \quad (5.9)$$

$$y = X_0, \quad (5.10)$$

where  $\dot{X}_0$  is the dynamical extension that allows us to integrate the noisy signal in order to recover a filtered signal, while  $\dot{X}_{n+1}$  allows us to put the unknown regulation function as a new state. Thus, the task becomes the estimation of this new state (a standard task for an observer), and therefore the function  $\Omega$  is related to the unknown dynamics of the new state. At this point,  $X \in \mathbb{R}^{n+2}$ , and furthermore the following equation is generated

$$\dot{X} = AX + B + E\delta ,$$

where  $AX$  is the linear part of the previous system such that  $A$  is a matrix equivalent in form to a Brunovsky matrix,  $B = [0, \dots, 0, \Omega(X)]^T$  and  $E = [1, 0, \dots, 0]^T$ .

We will need now the following result proven in Ref. [88].

*An asymptotic-type observer of the system  $\Xi$  is given as follows:*

$$\hat{\Xi}: \begin{cases} \dot{X}_0 = CX + \theta_1 (y_0 - \hat{y}_0) \\ \dot{X} = \hat{S} + \ell(\hat{X}) + \theta_2 (y_0 - \hat{y}_0) \\ \dot{\hat{S}} = \theta_3 (y_0 - \hat{y}_0) \\ \hat{y}_0 = \hat{X}_0 , \end{cases} \quad (5.11)$$

where the gain vector  $\theta$  of the observer is given by

$$\theta = S_\theta^{-1} C^T , \quad (5.12)$$

$$S_{\theta;i,j} = \left( \frac{S_{i,j}}{\vartheta^{i+j+1}} \right) . \quad (5.13)$$

Each entry of the matrix  $S_\theta$  is given by the above equation, where  $S_\theta$  is a  $n \times n$  matrix ( $i$  and  $j$  run from 1 to  $n$ ), and  $S_{i,j}$  are entries of a symmetric positive definite matrix that do not depend on  $\vartheta$ . Thus,  $S_{i,j}$  are such that  $S_\theta$  is a positive solution of the algebraic Riccati equation

$$S_\theta \left( A + \frac{\vartheta}{2} I \right) + \left( A + \frac{\vartheta}{2} I \right) S_\theta = C^T C . \quad (5.14)$$

In all formulas,  $C = [1, 0, \dots, 0]$ . In the multivariable case we must create one matrix  $S_\theta$  for each block corresponding to each output. It is worth mentioning that we can think about this observer as a ‘slave’ system that follows the ‘master’ system, which is precisely the real experimental system. In addition,  $S_\theta$ , as functional components of the gain vector, guarantees the accurate estimation of the observer through the convergence to zero of the error dynamics, i.e., the dynamics of the difference between the measured state and its corresponding estimated state. One can see that  $\vartheta$  generates an extra degree of freedom that can be tuned by the user such that the performance of the software sensor becomes satisfactory for him.

In [90] it has been shown that such an observer has an exponential-type decay for any initial conditions. Notice that a dynamic extension is generated by considering the measured output of

the original system as new additional dynamics with the aim to filter the noise. This procedure eliminates most of the noise in the new output of the system. The reason of the filtering effect is that the dynamic extension acts at the level of the observer as an integration of the output of the original system, (see the first equation of the system  $\Xi$  and the error part in the equations of system  $\hat{\Xi}$ ). The integration has averaging effects upon the noisy measured states. More exactly, the difference between the integral of the output of the slave part of system  $\hat{\Xi}$  and the integral of the output of the original system gives the error and the observer is planned in such a way that the error dynamics goes asymptotically to zero, which results in the recovering of both the filtered state and the unmeasured states.

### Particular Case

For gene regulation processes, which are of interest to us here, we merely apply the aforewritten system of equations corresponding to the asymptotic observer  $\hat{\Xi}$

$$\dot{X}_1 = K_{1,3}r(X_3) - \gamma_1 X_1 \quad (5.15)$$

$$\dot{X}_2 = K_{2,1}X_1 - \gamma_2 X_2 \quad (5.16)$$

$$\dot{X}_3 = K_{3,2}X_2 - \gamma_3 X_3 . \quad (5.17)$$

The pictorial representation of this system of equations is given in Fig. 5.2.

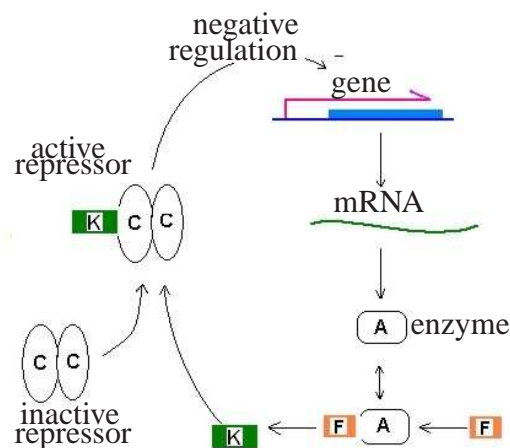


Figure 5.2: The genetic regulatory system given by Eqs. (5.15) - (5.17) involving end-product inhibition according to De Jong [73]. A is an enzyme and C a repressor protein, while K and F are metabolites. The mathematical model, as used by De Jong and by us, takes into account experiments where only metabolite K is measured.

The values of the parameters given in Table 1, without necessarily being the experimental values, are however consistent with the requirements of the model.

Table 5.1: Parameters of the model

Symbol	Meaning	Value (arb. units)
$K_{1,3}$	Production constant of mRNA	0.001
$K_{2,1}$	Production constant of protein A	1.0
$K_{3,2}$	Production constant of metabolite K	1.0
$\gamma_1$	Degradation constant of mRNA	0.1
$\gamma_2$	Degradation constant of protein A	1.0
$\gamma_3$	Degradation constant of metabolite K	1.0
$\vartheta$	Hill's threshold parameter	1.0

Using the structure given by the equations of  $\hat{\Xi}$ , the explicit form of the software sensor is:

$$\dot{\hat{X}}_0 = \hat{X}_3 + \theta_1(y_0 - \hat{X}_3) \quad (5.18)$$

$$\dot{\hat{X}}_1 = X_4 - \gamma_1 X_1 + \theta_2(y_0 - \hat{y}_0) \quad (5.19)$$

$$\dot{\hat{X}}_2 = K_{2,1} X_1 - \gamma_2 X_2 + \theta_3(y_0 - \hat{y}_0) \quad (5.20)$$

$$\dot{\hat{X}}_3 = K_{3,2} X_2 - \gamma_3 X_3 + \theta_4(y_0 - \hat{y}_0) \quad (5.21)$$

$$\dot{\hat{X}}_4 = \theta_5(y_0 - \hat{X}_3), \quad (5.22)$$

$$\hat{y}_0 = \hat{X}_0. \quad (5.23)$$

Notice that this dynamic structure does not involve the regulation function.

We can solve Eq. (5.14) and for numerical purposes we choose  $\vartheta = 2.5$  and the standard deviation of the Gaussian noise of 0.001. Figure 5.3 shows the numerical simulation that illustrates the filtering effect of the software sensor over the noisy measured state.

On the other hand, Fig. 5.4 shows the results of a numerical simulation, where the solid lines stand for the true states and the dotted lines indicate the estimates, respectively.



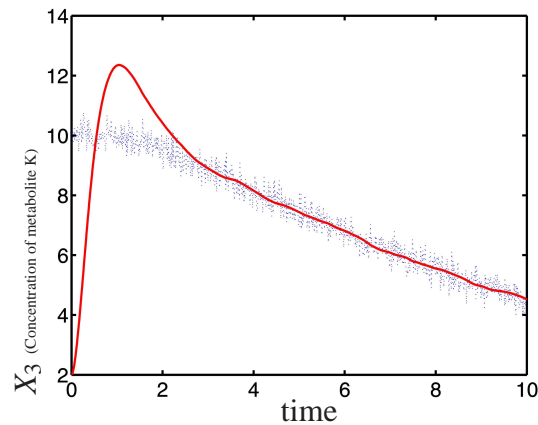


Figure 5.3: Numerical simulation: solid lines represent the filtered states and the dotted lines represent the noisy measured state for the evolution in time of metabolite K concentration. Notice that the initial bad estimation is due to the initial conditions that have been chosen far away from the real ones. This behaviour could be improved with a better knowledge of the initial conditions. The units of the two axes are arbitrary, i.e., the model is nondimensional.

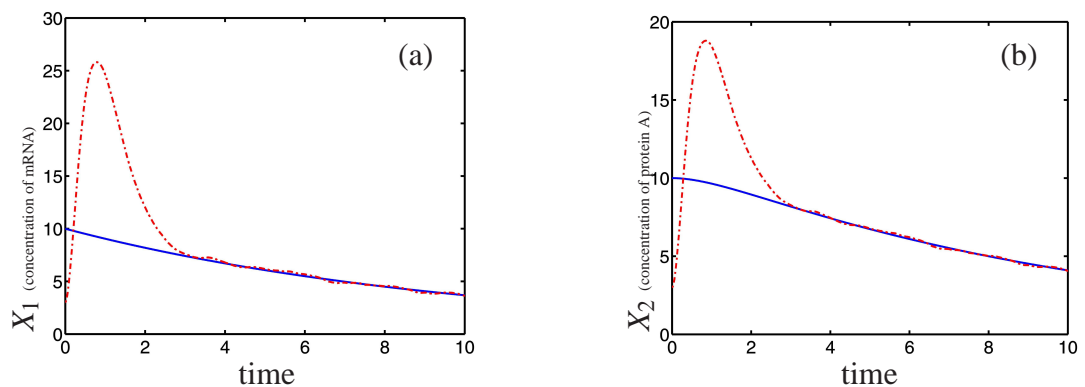


Figure 5.4: Numerical simulation: solid lines represent the true states generated by the original process endowed with the Hill regulatory function and dotted lines represent the estimated concentrations provided by the software sensor without any knowledge about the regulatory function. Plot (a) represents the evolution of mRNA concentration in time and plot (b) the variation of the concentration of protein A in time. The two axes have arbitrary units.

## 5.6 Three-Gene Circuit Case

In this section we extend the previous results to a more complicated case that can occur in prokaryotic cells. We study a more elaborated system where one regulator affects different

promoters and transcription units. The case corresponds to the coupled regulation of three genes in which the metabolite resulting from the translation of gene 1 becomes the substrate for the synthesis of the metabolite catalyzed by the enzyme translated from gene 2, and similarly for gene 3, but the metabolite 3 becomes the repressor of all the three genes involved, as shown in Fig. (5.5).

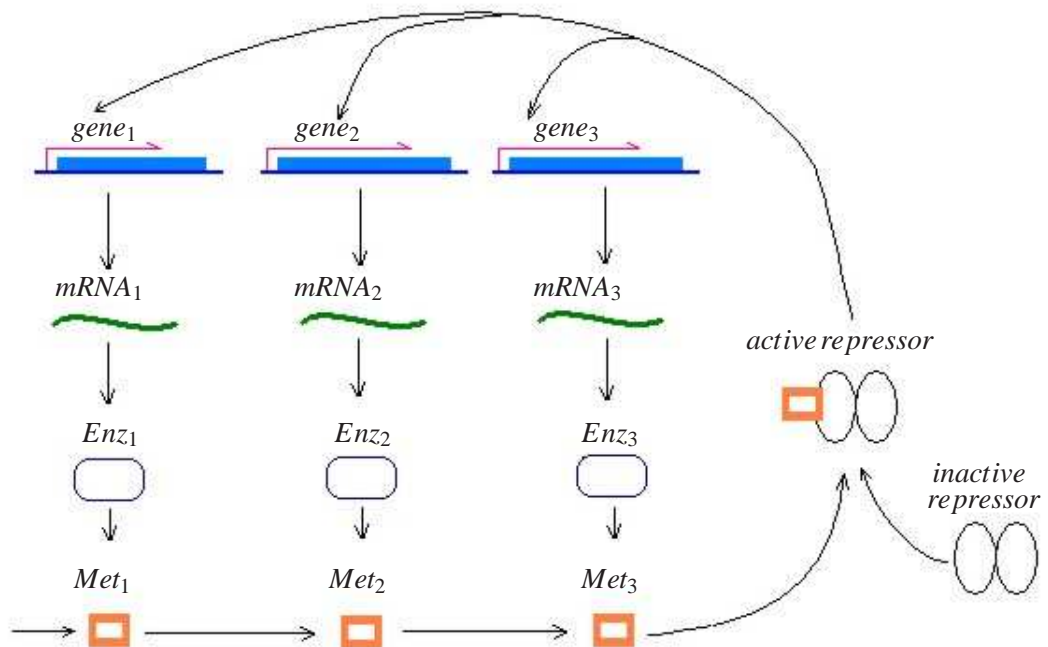


Figure 5.5: The three-gene regulatory circuit under consideration.

In this case the model is given by an extension of the model given by Eqs. (5.1,5.2). That results

in the following system of differential equations:

$$\frac{d}{dt}[mRNA_1] = K_1R([Met_3]) - \gamma_1[mRNA_1] \quad (5.24)$$

$$\frac{d}{dt}[Enz_1] = K_2[mRNA_1] - \gamma_2[Enz_1] \quad (5.25)$$

$$\frac{d}{dt}[Met_1] = K_3[Enz_1] - \gamma_3[Met_1] - \alpha_1[Enz_2] \quad (5.26)$$

$$\frac{d}{dt}[mRNA_2] = K_4R([Met_3]) - \gamma_4[mRNA_2] \quad (5.27)$$

$$\frac{d}{dt}[Enz_2] = K_5[mRNA_2] - \gamma_5[Enz_2] \quad (5.28)$$

$$\frac{d}{dt}[Met_2] = K_6[Enz_2] - \gamma_6[Met_2] - \alpha_2[Enz_3] \quad (5.29)$$

$$\frac{d}{dt}[mRNA_3] = K_7R([Met_3]) - \gamma_7[mRNA_3] \quad (5.30)$$

$$\frac{d}{dt}[Enz_3] = K_8[mRNA_3] - \gamma_8[Enz_3] \quad (5.31)$$

$$\frac{d}{dt}[Met_3] = K_9[Enz_3] - \gamma_9[Met_3] , \quad (5.32)$$

where  $[mRNA_i]$ ,  $[Enz_i]$  and  $[Met_i]$  represent the concentration of mRNA, enzymes and metabolites for each gene respectively. We select as the measured variables the metabolites because we want to show that through the measurement of stable molecules such as the metabolites, it is possible to infer the concentration of unstable molecules such as the mRNAs. Note that the equations are coupled through the dynamics of the metabolites. Moreover, we will assume that the dynamics of mRNA is bounded but unknown.

As we showed in the previous sections our new system can be written as:

$$\dot{X}_1 = X_2 + d_1 \quad (5.33)$$

$$\dot{X}_2 = K_3 X_3 - \gamma_3 X_2 - \alpha_1 X_8 \quad (5.34)$$

$$\dot{X}_3 = K_2 X_4 - \gamma_2 X_3 \quad (5.35)$$

$$\dot{X}_4 = X_5 \quad (5.36)$$

$$\dot{X}_5 = \phi_1(X) \quad (5.37)$$

$$\dot{X}_6 = X_7 + d_2 \quad (5.38)$$

$$\dot{X}_7 = K_6 X_8 - \gamma_6 X_7 - \alpha_2 X_{13} \quad (5.39)$$

$$\dot{X}_8 = K_5 X_9 - \gamma_5 X_8 \quad (5.40)$$

$$\dot{X}_9 = X_{10} \quad (5.41)$$

$$\dot{X}_{10} = \phi_2(X) \quad (5.42)$$

$$\dot{X}_{11} = X_{12} + d_3 \quad (5.43)$$

$$\dot{X}_{12} = K_9 X_{13} - \gamma_9 X_{12} \quad (5.44)$$

$$\dot{X}_{13} = K_8 X_{14} - \gamma_8 X_{13} \quad (5.45)$$

$$\dot{X}_{14} = X_{15} \quad (5.46)$$

$$\dot{X}_{15} = \phi_3(X), \quad (5.47)$$

where  $mRNA_1 = \dot{X}_4$ ,  $mRNA_2 = \dot{X}_9$ ,  $mRNA_3 = \dot{X}_{14}$ ,  $Enz_1 = \dot{X}_3$ ,  $Enz_2 = \dot{X}_8$ ,  $Enz_3 = \dot{X}_{13}$ ,  $Met_1 = \dot{X}_2$ ,  $Met_2 = \dot{X}_7$ ,  $Met_3 = \dot{X}_{12}$ ,  $d_i$  represent the noise,  $\phi_i(X)$  stand for the unknown dynamics. In addition, the previous systems can be written in the matricial forma as:

$$\dot{X} = \bar{A}X + \bar{B}(X) + Ed, \quad X \in \mathbb{R}^n \quad (5.48)$$

$$y = \bar{C}X = (C_1 X^1 \dots C_m X^m)^T, \quad (5.49)$$

where in this case  $X^i \in \mathbb{R}^{\lambda_i}$  is the  $i$ th partition of the state  $X$  so that  $X = [(X^1)^T, \dots, (X^m)^T]^T$  and  $\sum_{i=1}^m \lambda_i = n$ ;  $\bar{A} = \text{diag}[A^1, \dots, A^m]$  where  $A^i$  is  $\lambda_i \times \lambda_i$  such that  $S_\theta^i$  in the equation (5.14) is invertible;  $C = \text{diag}[C_1, \dots, C_m]$ , where  $C_i = [1, 0, \dots, 0] \in \mathbb{R}^{\lambda_i}$ ;  $B(X) = \text{diag}[B^1(X)^T, \dots, B^m(X)^T]^T$ ;  $E = \text{diag}[E_1, \dots, E_m]$ , where  $E_i = [1, 0, \dots, 0] \in \mathbb{R}^{\lambda_i}$ .

According to the scheme presented in the previous section we construct an observer through the following system of differential equations

$$\dot{\hat{X}}_1 = \hat{X}_2 + \theta_{11}(X_1 - \hat{X}_1) \quad (5.50)$$

$$\dot{\hat{X}}_2 = K_3\hat{X}_3 - \gamma_3\hat{X}_2 - \alpha_1\hat{X}_8 + \theta_{12}(X_1 - \hat{X}_1) \quad (5.51)$$

$$\dot{\hat{X}}_3 = K_2\hat{X}_4 - \gamma_2\hat{X}_3 + \theta_{13}(X_1 - \hat{X}_1) \quad (5.52)$$

$$\dot{\hat{X}}_4 = \hat{X}_5 + \theta_{14}(X_1 - \hat{X}_1) \quad (5.53)$$

$$\dot{\hat{X}}_5 = \theta_{15}(X_1 - \hat{X}_1) \quad (5.54)$$

$$\dot{\hat{X}}_6 = \hat{X}_7 + \theta_{21}(X_6 - \hat{X}_6) \quad (5.55)$$

$$\dot{\hat{X}}_7 = K_6\hat{X}_8 - \gamma_6\hat{X}_7 - \alpha_2\hat{X}_{13} + \theta_{22}(X_6 - \hat{X}_6) \quad (5.56)$$

$$\dot{\hat{X}}_8 = K_5\hat{X}_9 - \gamma_5\hat{X}_8 + \theta_{23}(X_6 - \hat{X}_6) \quad (5.57)$$

$$\dot{\hat{X}}_9 = \hat{X}_{10} + \theta_{24}(X_6 - \hat{X}_6) \quad (5.58)$$

$$\dot{\hat{X}}_{10} = +\theta_{25}(X_6 - \hat{X}_6) \quad (5.59)$$

$$\dot{\hat{X}}_{11} = \hat{X}_{12} + \theta\gamma_9\hat{X}_{12} + \theta_{32}(X_{11} - \hat{X}_{11}) \quad (5.60)$$

$$\dot{\hat{X}}_{13} = K_8\hat{X}_{14} - \gamma_8\hat{X}_{13} + \theta_{33}(X_{11} - \hat{X}_{11}) \quad (5.61)$$

$$\dot{\hat{X}}_{14} = \hat{X}_{15} + \theta_{34}(X_1 - \hat{X}_1) \quad (5.62)$$

$$\dot{\hat{X}}_{15} = \theta_{35}(X_{11} - \hat{X}_{11}), \quad (5.63)$$

where  $\theta_i$  stand for the observer gain values. Note, that this extension is not a direct application of that developed by Aguilar et al. [88] in the sense that this is a extension to the multivariable case. In addition, the matrix  $A_i$  is equivalent to a matrix of Brunovsky form, which guarantees the existence, uniqueness and invertibility of the matrix solution  $S_{\theta^i}$  [91]. (The existence and the uniqueness of  $S_{\theta^i}$  follows from the facts that  $-\frac{\theta_i}{2}I - A_i$  is of Hurwitz-type and that the pair  $(-\frac{\theta_i}{2}I - A_i, C_i)$  is observable [92]).

Figure 5.6 shows the numerical simulation of the filtering effect of the software sensor over the noisy measured state in this case. On the other hand, Fig. 5.7 displays the results of a numerical simulation of the true states (solid lines) and the estimates (dotted lines).

## 5.7 Conclusion

In this research, a simple software sensor was designed for a schematic gene regulation dynamic process involving end-product inhibition in single gene, operon and three gene circuit cases. This sensor effectively rebuilds the unmeasured concentrations of mRNA and the corresponding enzyme. Thus, the limitation of those experiments in which only the concentration of the catalytically synthesized metabolite is available, can be overcome by employing the simple software sensor applied here. This is a quite natural case if one takes into account that

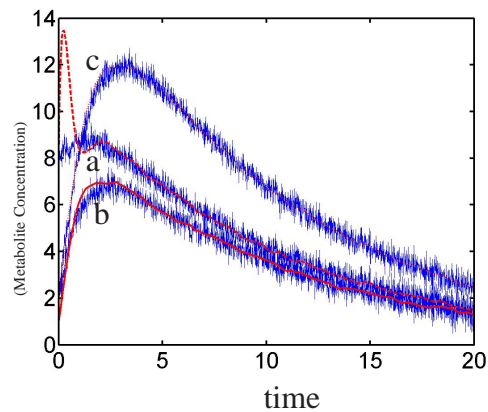


Figure 5.6: Numerical simulation: solid lines represent the filtered states obtained from the noisy measured states for the evolution in time of metabolite concentrations, where a, b and c correspond to metabolite 1, 2, and 3, respectively. The units of the two axes are arbitrary (nondimensional model).

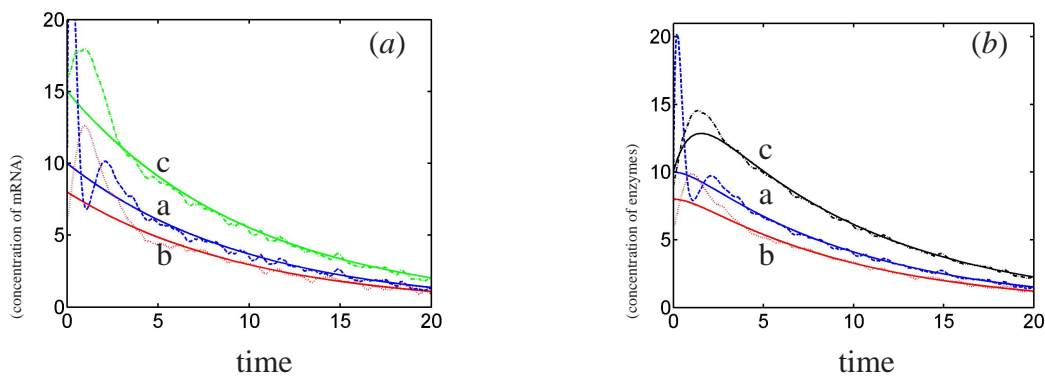


Figure 5.7: Numerical simulation: solid lines represent the true states generated by the original process endowed with the Hill regulatory function and dotted lines represent the estimated concentrations provided by the software sensor without any knowledge about the regulatory function; a, b and c correspond to molecule 1, 2 and 3, respectively. Plot (a) represents the evolution of  $mRNA_i$  concentrations in time and plot (b) the variation of the concentration of the corresponding enzymes in time. The axes of the graph have arbitrary units.

metabolites are quite stable at the molecular level. At the same time, we can reproduce the concentrations of the unstable molecules of mRNA. This is a difficult task in experiments, despite the fact that the mRNA dynamics has been partially or even totally unspecified.

The same scheme philosophy to build the observer is applied to a three-gene circuit with the

purpose to show that the software sensor concept could be in usage in a forward engineering approach. In this research however, we mentioned that we were able to show that the observer scheme designed in [88] for the single output case works well also in a multiple variable case as embodied by a particular genetic circuit given in Fig. (5.5). The most stringent mathematical requirement for this extended applicability to the multiple output case is described below. The linear part of the dynamic system should be a matrix by blocks in which each of the blocks should be of Brunovsky equivalent form. In addition, each subsystem corresponding to a superior block depends only on the subsystem corresponding to the next nearest block. This is a feature similar to the property of Markoff processes. The Brunovsky equivalent form of the matrix blocks  $A_i$  together with the structure of the corresponding output vector  $C_i$  generate an observable pair  $(A_i, C_i)$ , giving us the capability to infer the internal states of the gene network through the knowledge of its external outputs. However, the special Brunovsky equivalent form of the blocks leads to the possible biological interpretation that each block of the linear part of the differential system represents only that contribution of the gene regulation mechanism that comes from reactions occurring in a cascade fashion.

Another important issue that we tackled in this work is related to the way of adding the noise to the output of the dynamic system. Even though this is a typical situation from the standpoint of control process theory, to the best of our knowledge it has not yet been applied in the biological context of gene regulation processes. We stress that this way of including noise effects could have both intrinsic and extrinsic interpretations and therefore assure a more general approach of the noise problems. For example, in phenomenological terms, perturbations on the cells due to the measuring devices and the experimental conditions, together with the noise produced by the nature of the electronic instrumentation, could be equally described in this way.

In addition, this type of nonlinear observer could be used as an online filter being robust with respect to model uncertainties, i.e., neither a known regulation function nor the parameter  $K_{1,3}$  is required.

**Part I**  
**Bibliography**



---

---

## Bibliography

---

- [1] Ibarra-Junquera, V., and Rosu, H.C., (2006), *PI-Controlled Bioreactor as a Generalized Liénard System*. Accepted in *Computers and Chemical Engineering*, (arXiv:nlin.CD/0502049).
- [2] Lo, S.N., & Cholette, A., (1983), *Multiplicity of a conversion in cascade in imperfectly stirred tank reactors*. *Chem. Eng. Sci.* **38**, 367.
- [3] Liou, C.T., & Y.S. Chien, (1991), *The effect of nonideal mixing on input multiplicity in a CSTR*. *Chem. Eng. Sci.* **46**, 2113.
- [4] Sree, R.P., & Chidambaram, M.A., (2002), *Identification of unstable transfer model with a zero by optimization method*. *J. Indian Inst. Sci.* **82**, 219.
- [5] Sree, R.P., & Chidambaram, M.A., (2003), *Control of unstable bioreactor with dominant unstable zero*. *Chem. Biochem. Eng. Q.* **17**, 139.
- [6] Sree, R.P., & Chidambaram, M.A., (2003), *A simple method of tuning PI controllers for unstable systems with a zero*. *Chem. Biochem. Eng. Q.* **17**, 207.
- [7] Kumar, V.R., & Kulkarni, B.D., (1994), *On the operation of a bistable CSTR: a strategy employing stochastic resonance*. *Chem. Eng. Sci.* **49**, 2709.
- [8] Gaiko, V.A., (2000), *Global bifurcations of limit cycles*.  
<http://www.neva.ru/journal>, *Diff. Eqs. and Control Processes*, No. 3, 1.
- [9] Gordillo, F., Salas, F., Ortega, R., Aracil, J., (2002), *Hopf bifurcation in indirect field-oriented control of induction motors*. *Automatica* **38**, 829.
- [10] Ogata, K., (2001), *Modern Control Engineering*, 4th edition, Prentice Hall.

- [11] Xiao D., & Zhang, Z., (2000), *On the uniqueness and nonexistence of limit cycles for predator-prey systems*. Nonlinearity **16**, 1.
- [12] Li, W., Llibre, J., & Zhang, X., (2004), *Melnikov functions for period annulus, nondegenerate centers, heteroclinic and monoclinc cycles*. Pacific J. Math. **213**, 49.
- [13] Giacomini, H., & Neukirch, S., (1997), *Number of limit cycles of the Liénard equation*. Phys. Rev. E **56**, 3809.
- [14] Giacomini, H., & Neukirch, S., (1997), *Improving a method for the study of limit cycles of the Liénard equation*. Phys. Rev. E **57**, 6573.
- [15] Giacomini, H., & Neukirch, S., (1998), *Algebraic approximations to bifurcation curves of limit cycles for the Liénard equation*. Phys. Lett. A **244**, 53.
- [16] Du, Z., (2004), *On the critical periods of Liénard systems with cubic restoring forces*. Int. J. Math. and Math. Sci. **61**, 3259.
- [17] Cristopher, C.J., & Devlin, J., (1997), *Isochronous center in planar polynomial systems*. SIAM J. Math. Anal. **28**, 162.
- [18] Albarakati, W.A., Lloyd, N.G., & Pearson J.M., (2000), *Transformation to Liénard form*. Electronic J. Diff. Eqs., Vol. 2000, No. 76, 1.
- [19] Kalman, R. E., (1960), *Contributions to the theory of optimal control*. Bull. Soc. Mat. Mexicana, 5, 102-119.
- [20] Weigand, W. A., (1981), *Maximum cell productivity by repeated fed-batch culture for constant yield case*. Biotechnology and Bioengineering Volume 23, Issue 2 , 249-266.
- [21] Bourdache-Siguerdidjane, H., and Fliess, M., (1987), *Optimal Feedback Control of Nonlinear Systems*. Automatica, vol. 23, No. 3, 365-372.
- [22] Palanki, S., Kravaris, C., and Wang, H.Y., (1993), *Synthesis of State Feedback Laws for End-Point Optimization in Batch Processes*. Chem. Eng. Sci., Vol. 48, No. 1, 135-152.
- [23] Van Impe, J.F., Bastin, G.P., (1993) *Optimal adaptive control of fed-bacht fermentation processes with multiple substrates*. J. of Process Ctrl., Vol. 14(7), 795-805.

- [24] Palanki, S., and Rahman, A.K.M.S., (1994), *State feedback synthesis for on-line optimization of batch reactors with a non-affine manipulated input and free terminal time*. Proceedings of the American Control Conference, Baltimore, Maryland, June, 3076-3079.
- [25] Palanki, S., Feteih, S., and Sadhukhan, D., (1995), *On-line optimization of batch processes using neural networks coupled with optimal state feedback*. Proceedings of the American Control Conference, Seattle, Washington, June, 1772-1776.
- [26] Van Impe, J.F., Bastin, G.P., (1995), *Optimal adaptive control of fed-batch fermentation processes*. J. of Process Ctrl., Vol. 14(7), 795-805.
- [27] Rahman, A.K.M.S., and Palanki, S., (1996), *State feedback synthesis for on-line optimization of batch reactors with nonlinear manipulated input*. Chem. Eng. Sci., Vol. 51, No. 3, 449-459.
- [28] Rahman, A.K.M.S., and Palanki, S., (1996), *On-line optimization of batch reactors with two manipulated inputs*. Computers Chem. Engng. Vol. 20, Suppl., S1023-S1028.
- [29] Da Costa, A.C., Lima, E.L. and Alves, T. L.M., (1997), *Start-up Strategy for Continuous Bioreactors*. Braz. J. Chem. Eng. vol.14, no.2.
- [30] Rahman, A.K.M.S., and Palanki, S., (1998), *State feedback synthesis for on-line optimization of batch reactors with multiple manipulated inputs*. Computers Chem. Engng. Vol. 10, 1429-1439.
- [31] Shukla, P.K., and Pushpavanam, S., (1998), *Optimization of biochemical reactor -an analysis of different approximations of fed-batch operation*. Chem. Eng. Sci., Vol. 53, No. 2, 341-352.
- [32] Visser, E., Srinivasan, B., Palanki, S., and Bonvin, D., (2000), *A feedback-based implementation scheme for batch process optimization*. J. Process Ctrl., 10, 399-410.
- [33] Betlem B. H. L., Mulder P. and Roffel B., (2002), *Optimal mode of operation for biomass production*. Chem. Eng. Sci., Vol. 57(14), 2799-2809.

- [34] Palanki, S., and Vemuri, J., (2003), *End-point optimization of batch chemical processes*. IEEE Conf. on Decision and Control, Maui, Hawaii, December 9-12.
- [35] Srinivasan, B., Palanki, S., and Bonvin, D., (2003), *Dynamic Optimization of Batch Processes: I Characterization of the Nominal Solution*. *Comp. Chem. Eng.*, 27, 1, 1-26.
- [36] Srinivasan, B., Palanki, S., and Bonvin, D., (2003), *Dynamic Optimization of Batch Processes: II Role of Measurements in Handling Uncertainty*. *Comp. Chem. Eng.*, 27, 1, 27-44.
- [37] Smets, I.Y., Claes, J.E., November, E.J., Bastin, G.P., Van Impe, J.F., (2004), *Optimal adaptive control of (bio)chemical reactors: past, present and future*. *J. of Process Ctrl.*, Vol. 14(7), 795-805.
- [38] Isidori, A., (1995), *Nonlinear Control Systems*. (Springer N.Y., Third Edition).
- [39] Szerderkényi, G., Kovács, M., and Hangos, K.M., (2002), *Reachability of nonlinear fed-batch fermentation processes*. *Int. J. Robust nonlinear control*. 12, 1109-1124, (DOI: 10.1002/rnc.686).
- [40] Szerderkényi, G., Kristensen, N.R., Hangos, K.M., & Jørgensen, S.B., (2002), *Nonlinear analysis and control of a continuous fermentation process*. *Computers Chem. Engng*. Vol. 26, 659-670.
- [41] Ibarra-Junquera, V., Femat, R., & Lizárraga, D.A., (2004), *On Structure of a Bioreactor for Cell Producing: Effects by Inhibitory Kinetic.*, 2nd IFAC SYMPOSIUM on SYSTEM, STRUCTURE and CONTROL., Oaxaca, México, (2004).
- [42] Byrnes, C., Isidori, A., & Willems, J.C., (1991), *Passivity, Feedback Equivalence, and the Global Stabilization of Minimum Phase Nonlinear Systems.*, *IEEE Trans. Automat. Contr.*, vol.36,(11) 1228-1240.
- [43] Brown, R. & Kocarev, L., (2000), *A unifying definition of synchronization for dynamical systems*. *Chaos*, Vol. 10, No. 2., 344-349.

- [44] Josić, K., (1998), *Invariant Manifolds and Synchronization of Coupled Dynamical Systems*. Phys. Rev. Letters., Vol. 80, No. 14., 3053-3056.
- [45] Jakubczyk, B. and Respondek, W. (1980), *On linearization of control systems*. Bull. Acad. Pol. Sci. Ser. Sci. Math. 18: 517-522.
- [46] Wang, G.B., Peng, S.S., and Huang, H.P., (1990), *A sliding observer for nonlinear process control*. Chem. Eng. Science. vol. 52, 5: 787-805.
- [47] Behtash S. (1990), *Robust output tracking for non-linear systems*. Int J. Control. vol. 51, 6: 1391-1407.
- [48] Ibarra-Junquera, V., Escalante-Minakata, P., Murguía-Ibarra, J.S., & Rosu, H.C., (2006), *Inferring mixed-culture growth from total biomass data in a wavelet approach*. Accepted in Physica A ( arXiv.org:physics/0512186).
- [49] Szambelan, K., Nowak, J., & Czarnecki, Z., (2004), *Use of Zymomonas mobilis and Saccharomyces cerevisiae mixed with Kluyveromyces fragilis for improved ethanol production from Jerusalem artichoke tubers*. Biotech. Letters 26: 845-848.
- [50] Fleet, G.H., (2003), *Yeast interactions and wine flavour*. Int. J. of Food Microbiol., 86: 11-22.
- [51] de Souza Liberal, A.T., da Silva Filho, E.A., de Moraes, J.O.F., Simões, D.A., (2005), *Contaminated yeast detection in industrial ethanol fermentation must by rDNA-PCR*. Letters in Applied Microbiol., **40**, 19-23.
- [52] Granchi, L., Bosco, M., Messini, A., Vicenzini, M., (1999), *Rapid detection and quantification of yeast species during spontaneous wine fermentation by PCR-RFLP analysis of rDNA ITS region*. J. of Applied Microbiol., **87**, 949-956.
- [53] Callister, S.J., Ayala-del-Rio, H.L. and Hashsham, S.A., (2003), *Quantification of a single population in a mixed microbial community using a laser integrated microarray scanner*. J. Environ. Eng. Sci./Rev. gen. sci. env., **2**, 247-253.

- [54] Madrid, R. & Felice, C., (2005), *Microbial Biomass Estimation*. Critical Reviews in Biotechnology, **25**, 97-112(16).
- [55] Pons, M.N., Vivier, H., Rémy, J.F., Dodds, J.A., (1993), *Morphological characterization of yeast by image analysis*. Biotechnol. and Bioeng. Vol. 42, Issue 11,1352-1359.
- [56] Reeves, G.T., Narang, A., & Pilyugin S.S., (2004), *Growth of mixed cultures on mixtures of substitutable substrates: the operating diagram for a structured model*. Journal of Theoretical Biology, **226**, 142-157. (Available online at [www.sciencedirect.com](http://www.sciencedirect.com))
- [57] Muzy, J.F., Bacry, E. and Arneodo, A.,(1994), *The multifractal formalism revisited with wavelets*. Int. Journal of Bifurcations and Chaos, **4**, 245-302.
- [58] Arnéodo, A., Bacry, E., Jaffard, S. and Muzy, J.F., (1997), *Oscillating singularities on Cantor sets. A grand - canonical multifractal formalism*. J. Stat. Phys., **87**, 179-209.
- [59] Arnéodo, A., Bacry, E., Jaffard, S. and Muzy, J.F., (1998), *Singularity spectrum of multifractal functions involving oscillating singularities*. J. Fourier Analysis and Applications, **4**, 159-174.
- [60] Daubechies, I., (1992), *Ten lectures on wavelets*. SIAM, Philadelphia, PA, 1992.
- [61] Mallat, S. and Hwang, W.L., (1992), *Singularity detection and processing with wavelets*. IEEE Trans. Inform. Theory, **38**, 617–643.
- [62] Mallat, S., (1999), *A Wavelet Tour of Signal Processing*, 2nd. Edition, Academic Press, Inc., New York, 1999; See also: Bacry, E., Muzy, J.F. and Arnéodo, A., (1993), *Singularity spectrum of fractal signals from wavelet analysis: Exact results*. J. Stat. Phys. **70**, 635–674.
- [63] Herrmann, F.J.,(1997), *A scaling medium representation, a discussion on well-logs, fractals and waves*. Ph.D. Thesis, Delft University of Technology, Netherlands, 1997.
- [64] Ibarra-Junquera, V., Torres, L.A., Rosu, H.C., Argüello, G., and Collado-Vides, J., (2005), *Nonlinear software sensor for monitoring genetic regulation processes with noise and modeling errors*, Phys. Rev. E **72**, 011919.

- [65] Lewin, B., *Genes VII*, (Oxford Univ. Press, Oxford, 1999).
- [66] See <http://www.iee.org/sb>
- [67] Yi T.-M., Huang Y., Simon M.I. , and Doyle J., (2000), *Robust perfect adaptation in bacterial chemotaxis through integral feedback control*. Proc. Nat. Ac. Sci. USA **97**, 4649 .
- [68] Barkai N., & Leibler, S., (1997), *Robustness in simple biochemical networks*. Nature **387**, 913; Alon, U., Surette, M.G., Barkai, N., & Leibler, S., (1999), *Robustness in bacterial chemotaxis*. *ibid.* **397**, 168.
- [69] Luenberger D., (1966), *Observers for multivariable systems*. IEEE Trans. Autom. Control **11**, 190.
- [70] Hasty, J., Pradines, J., Dolnik, M., & Collins, J.J., (2000), *Noise-based switches and amplifiers for gene expression*. Proc. Nat. Acad. Sci. USA **97**, 2075; Kepler T., & Elston, T., (2001), *Stochasticity in transcriptional regulation: origins, consequences, and mathematical representations*. Biophys. J. **81**, 3116 (2001); Swain, P.S., Elowitz, M.B., & Siggia, E.D., (2002), *Intrinsic and extrinsic contributions to stochasticity in gene expression*. Proc. Nat. Acad. Sci. USA **99**, 12795; Blake, W.J., Kaern, M., Cantor, C.R., & Collins, J.J., (2003), *Noise in eukariotic gene expression*. Nature **422**, 633 (2003); Isaacs, F.J., Hasty, J., Cantor, C.R., & Collins, J.J., (2003), *Prediction and measurement of an autoregulatory genetic module*. Proc. Nat. Acad. Sci. USA **100**, 7714.
- [71] Davidson, E.H., Rast, J.P., Oliveri, P., Ransick, A., Calestani, C., Yuh, C.-H., Minokawa, T., Amore, G., Hinman, V., Arenas-Mena, C., Otim, O., Brown, C. T., Livi, C.B., Lee, P.Y., Revilla, R., Rust, A. G., Pan, Z.j., Schilstra, M.J., Clarke, P. J. C., Arnone, M. I., Rowen, L., Cameron, R. A., McClay, D.R., Hood, L., & Bolouri, H., (2002), *A genomic regulatory network for development*. Science **295**, 1669.
- [72] Lee, T.I., Rinaldi, N.J., Robert, F., Odom, D. T., Bar-Joseph, Z., Gerber, G. K., Hannett, N. M., Harbison, C. T., Thompson, C. M., Simon, I., Zeitlinger, J., Jennings, E. G., Murray, H. L., Gordon, D.B., Ren, B., Wyrick, J. J., Tagne, J.-B., Volkert, T. L., Fraenkel, E., Gifford, D. K., & Young, R. A., (2002), *Transcriptional regulatory networks in S.c.*. Science **298**, 799.
- [73] De Jong H. , (2002), *Modeling and Simulation of Genetic Regulatory Systems: A Literature Review*. J. Comp. Biol. **9**, 67.
- [74] Ozbudak, E., Thattai, M., Kurtser, I., Grossman, A.D., & van Oudenaarden, A., (2002), *Regulation of noise in the expression of a single gene*. Nature Genetics **31**, 69 (2002).
- [75] Brown, P.A., & Botstein, D., (1999), *Exploring the new world of the genome with DNA microarrays*. Nature Genetics **21** (suppl.), 33.

- [76] Lipshutz, R.J., Fodor, S.P., Gingeras, T.R., & Lockhart, D.J., (1999), *High density synthetic oligonucleotide arrays*. *Nature Genetics* **21** (suppl.), 20.
- [77] Lockhart, D.J., & Winzeler, E.A., (2000), *Genomic, gene expression and DNA arrays*. *Nature* **405**, 826.
- [78] Stephanopoulos, G., *Chemical Process Control*. (Prentice Hall, 1984).
- [79] Goodwin, B.C., *Temporal Oscillations in Cells*, (Academic Press, New York, 1963)
- [80] Tyson J.J., (1978), *The dynamics of feedback control circuits in biochemical pathways*. *Prog. Theor. Biology* **5**, 1, 1-62.
- [81] Nelson, D.L., & Cox, M.M., *Lehninger Principles of Biochemistry*, (Worth Publishers, New York, 2000) p. 1077.
- [82] Santillán, M., & Mackey, M.C., (2001), *Dynamic regulation of the tryptophan operon: A modeling study and comparison with experimental data*. *Proc. Nat. Acad. Sci. USA* **98**, 1364.
- [83] Wolovich, W.A., *Automatic Control Systems*, (Saunders College, Philadelphia, USA, 1994).
- [84] Krener A.J., & Isidori A., (1983), *Linearization by output injection and nonlinear observers*. *Systems Control Lett.* **3**, 47.
- [85] Gauthier, J.P., Hammouri, H., & Othaman, S., (1992), *A simple Observer for Nonlinear Systems Applications to Bioreactors*. *IEEE Trans. Aut. Ctrl.* **37**, 875.
- [86] Gauthier, J.P., & Bornard, G., (1981), *Obsevability for any  $u(t)$  of a class of Nonlinear Systems*. *IEEE Trans. Aut. Ctrl.* **AC-26**, 922, (4)922-926
- [87] Gauthier, J.P., & Kupka, I.A.K., (1994), *Observability and Observers for Nonlinear Systems*. *Siam J. Control and Optimization* **32**, 975, 975-994.
- [88] Aguilar R., Martínez-Guerra R., & Maya-Yescas R., (2003) *State estimation for a partially unknown nonlinear systems: a class of integral high gain observers*. *IEE Proc.:Contol Theory Appl.* **150**, 240, 975-994.
- [89] Aguilar-López R. , (2003), *Integral observers for uncertainty estimation in continuous chemical reactors: algebraic-differential approach*. *Chem. Eng. J.* **93**, 113, 975-994.
- [90] Martínez-Guerra R., Suarez R., & De León-Morales J., (2001), *Asymptotic tracking of a class of nonlinear systems by means of an observer*. *Int. J. Robust Nonlinear Control* **11**, 373, 233-244.



- 
- [91] Hermann R. & Krener A.J., (1977) *Nonlinear Controllability and Observability*. IEEE Trans. Aut. Ctrl. **AC-22**, 728,(6) 728-740.
- [92] Shim, H., Son, Y.I., & Seo, J.H., (2001), *Semi-global Observer for multi-output nonlinear systems*. Systems & Control Letters **41**, 233, 233-244.