DIME-UNIVERSITÀ DEGLI STUDI DI GENOVA
in cotutela con
CRAN-UNIVERSITÉ DE LORRAINE

**DOTTORATO DI RICERCA IN INGEGNERIA DELLE MACCHINE E DEI SISTEMI PER L'ENERGIA L'AMBIENTE ED I TRASPORTI**

**Curriculum Ingegneria Matematica & Simulazione**

**CICLO XXXIV**

---

# Modeling and Estimation of Biological Plants

---

**Presentata da: : Dyhia Bouhadjra**

**Coordinatore Dottorato**

**Prof. Roberto Cianci**

**Relatore**

**Prof. Angelo Alessandri**

**Prof. Ali Zemouche**

Esame finale anno 2022

## Acknowledgments

First and foremost, I am extremely grateful to my supervisors, professors Angelo Alessandri, and Patrizia Bagnerini from the University of Genoa in Italy and professor Ali Zemouche from the university of Lorraine in France, for their invaluable advice, continuous support, and patience during my Ph.D study. Their immense knowledge and plentiful experience have encouraged me in all the time of my academic research and daily life. Their continuous guidance and support, irrespectively of the clock index position, have catalyzed my development as a researcher, for which I am deeply grateful. Thank you for sharing your passion and (a small part of) the immense breadth of your knowledge with such generosity and kindness. Besides my supervisors, I would like to thank the rest of my thesis committee, professors Mauro Gaggero and Dalil Ichalal, who honored me by accepting to be the reviewers of my thesis. Their attentive reading and insightful comments helped me improve this manuscript and allowed me to widen my research from various perspectives. Thanks to everyone from the "Dipartimento di Ingegneria Meccanica, Energetica, Gestionale e dei Trasporti" at the University of Genoa. I especially want to thank Prof. Roberto Cianci for creating such a good Ph.D environment at DIME and always giving advice on Ph.D matters. I would like to express my gratitude to the people I met at IUT Longwy, especially, my fellow labmates and friends Hasni, Nada, Echrak, Hichem, and Shiv. Thanks for the stimulating discussions and for putting up with me through the stressful times, especially before deadlines. Thanks to my friends from inside and outside the university of Genoa especially, Chiara, Elena, and Seb. I know I can be difficult to deal with at the best of times but thank you for being there for me when I needed it and making me feel like I was home even if I was far from my family and my country. Last but not least, I would like to thank my parents, my sisters, and my brother for the tremendous support, hope, and love they had given to me, especially within the last three months of Ph.D. Without them, this day would not have been possible. Their moral paideia has constantly been an example to me for pursuing a life of virtue. I thank them, as well as the rest of my family for always believing in me.

*To the memory of my beloved sister Hayet.*

# Contents

# List of Figures

# List of Tables

# Notation and Acronyms

**Notation**

| | |
|---|---|
| $\mathbb{R}$ | Set of real numbers |
| $\mathbb{R}^+$ | Set of real positive numbers |
| $\mathbb{R}^n$ | Euclidean real space of dimension $n$ |
| $\mathbb{R}^{n\times m}$ | Set of real matrices of dimension $n \times m$ |
| $\|x\|$ | Euclidian norm of $x$, with $x \in \mathbb{R}^n$ |
| $\|A\|$ | Induced matrix norm defined as $\sup_{x\in\mathbb{R}^n}\left\{\frac{\lvert A\rvert}{\lvert Ax\rvert}\right\}$ |
| $A > 0$ | Symmetric positive definite matrix |
| $A \geq 0$ | Symmetric positive semi-definite matrix |
| $det(A)$ | Determinant of the matrix $A \in \mathbb{R}^{n\times n}$ |
| $\lambda_{min}(A)$ | Smallest eigenvalue of the matrix $A$ |
| $\lambda_{max}(A)$ | Largest eigenvalue of the matrix $A$ |
| $A^T$ | Transpose of the matrix $A$ |
| $A^{-1}$ | Inverse of the matrix $A$ |
| $I_n$ | An $n \times n$ identity matrix |
| $0_{n\times m}$ | Matrix of dimension $n \times m$ whose entries are all zeros |
| $diag(a_1, \ldots, a_n)$ | An $n \times m$ diagonal matrix with $a_i$ as its $i$-th diagonal element |

| | |
|---|---|
| $col(a_1, \ldots, a_n)$ | Column vector with elements $(a_1, \ldots, a_n)$ |

**Acronyms**

| | |
|---|---|
| DC | Dorsal Closure |
| EKF | Extended Kalman Filter |
| GRN | Genetic Regulatory Network |
| LMI | Linear Matrix Inequality |
| LPV | Linear Parameter Varying |
| LTI | Linear Time Invarying |
| LTV | Linear Time Varying |
| MHE | Moving Horizon Estimator |
| SI | Suscheptible Infected |

# List of publications

- **International Journals**

  1. A. Alessandri, F. Bedouhene, **D. Bouhadjra**, A. Zemouche, & P. Bagnerini (2021). Observer based Control for a Class of Hybrid Linear and Nonlinear Systems. *DCDS-S*, 14(4), 1213-1231.
  2. **D. Bouhadjra**, A. Alessandri, P. Bagnerini , & A. Zemouche (2022). Moving Horizon Estimation of Amnioserosa Cell Dynamics During Drosophila Dorsal Closure. *EJCON* (Accepted).
  3. **D. Bouhadjra**, & A. Zemouche (2022). New High-Gain Like Observer Design with Application to Gene Regulation and Epidemic Models. *IJACSP* (Submitted).

- **International Conferences**

  1. **D. Bouhadjra**, A. Zemouche, A. Alessandri, & P. Bagnerini (2020). High-Gain Nonlinear Observer using System State Augmentation. *Conference on Decision and Control, CDC'2020*, Jeju Island, Republic of Korea.
  2. **D. Bouhadjra**, A. Alessandri, P. Bagnerini, & A. Zemouche (2021). Modeling and Estimation of Amnioserosa Cell Mechanical Behavior using Moving Horizon Estimation. *American Control Conference, ACC'2021*, New Orleans, USA.
  3. H. Arezki, **D. Bouhadjra**, K. Chaib Draa, A. Zemouche, I. N'Doye & T. M. Laleg (2021). LMI Feasibility Improvement to Design Observers for a Class of Lipschitz Nonlinear Systems. *$60^{th}$ IEEE Conference on Decision and Control, CDC'2021*, Texas, USA.
  4. **D. Bouhadjra**, A. Alessandri, P. Bagnerini, & A. Zemouche (2021). A High-Gain Observer for Stage-Structured Susceptible-Infectious Epidemic Model with Linear Incidence Rate. *American Control Conference, ACC'2022*, Atlanta, USA.
  5. **D. Bouhadjra**, A. Alessandri, P. Bagnerini, & A. Zemouche (2021). A New High-Gain Observer for Estimation of mRNA and Protein Concentrations of a Genetic Regulatory Network. *European Control Conference, ECC'2022*, London, UK.

# Introduction

The knowledge of state variables of a dynamical system is crucial either to build a controller or simply to obtain real-time information on the system for decision-making or monitoring. One way to obtain such variables consists of combining a prior knowledge about physical systems with experimental data to design an algorithm, called observer, whose role is to process the incomplete and imperfect information to construct an estimate of the state variables.

The synthesis of such algorithms has attracted great attention from the automatic control community. Initially, the research was naturally oriented toward estimating the state of linear systems. An optimal solution was first developed in the early 1960s by Kalman [1] in a stochastic approach and later by Luenberger [2] in a deterministic context. These algorithms are still widely used nowadays but since linear systems cover a small percentage of processes, nonlinear solutions were quickly considered. A direct extension of these results to the nonlinear case is obtained by means of a local linearization of the system dynamics along the estimated trajectories: this is the principle of the extended Kalman filter [3]. However, the overall stability of this observer in the presence of strong nonlinearities has not been proven and its performance has been regularly challenged in practice.

Because of the lack of a general design method for nonlinear systems like in the linear case, several methods have been developed in the literature, where each method corresponds to a specific class of nonlinear systems. The concept of error linearization was introduced by [4], [5] for a certain class of mono-output nonlinear dynamical systems, then it was extended by [6], [7] to multi-output dynamic systems. From there, several algorithms have been proposed giving rise to various approaches: algebraic [8–10], geometric [11–13] and direct transformations [14].

A straightforward approach to nonlinear observer design is to use linear feedback. If the nonlinearities are globally Lipschitz, then it is possible to use LMIs (Linear Matrix Inequalities) combined with Lyapunov or Riccati equations which allow for elegantly treating the Lipschitz terms in the error dynamics [15–18]. More precisely, the gain of the observer is designed through the resolution of an LMI problem; consequently, an observer exists only if the considered LMI problem is feasible [15], [19]. As pointed out in [20], the feasibility of the LMI problems considered is generally not known a priori and is to be determined numerically. Several approaches have been explored in the literature to extend the class of systems considered. Among these, are sliding mode observers which are based on the theory of variable structure systems [21–24]. Their implementation difficulties have justified on the one hand the appearance of different variants concerning the choice of the mathematical function used and on the other hand an extension to higher orders.

Another important technique for nonlinear observer design is the high-gain observer introduced initially by [25] in 1973. The main idea is to use sufficiently large observer gains so as to dominate the nonlinearities (more precisely their Lipschitz constant) in the estimation error dynamics. The synthesis of these observers is carried out either in the initial coordinates (under certain structural hypotheses of the system) or in the canonical bases associated with uniformly observable or U-uniformly observable systems. Two key papers, published in 1992 [26,27], represent the beginning of two schools of research on high-gain observers. The work by Gauthier [26] started a line of work that is exemplified by [28–34]. This line of research covered a wide class of nonlinear systems and obtained global results under global growth conditions. Although high-gain observers are successfully applied in problems of estimation ( [28,35] or [32]), output feedback control [36–39] and output regulation (see [40] or [41]), their use in practical applications is made hard by a certain number of drawbacks related to numerical implementation. To overcome these drawbacks, several solutions have been advanced in the literature.

The so-called extended high-gain observer is presented in [42] which is composed of an Extended High Gain Observer (EHGO), for the estimation of the derivatives of the output, augmented with an Extended Kalman Filter (EKF) for the estimation of the states of the internal dynamics. Then, to account for the presence of disturbances acting on the system, several methods have been proposed based on gain adaptation methods [43–47]. The selection of a high gain stems also from the need to account for the nonlinearities in the error dynamics, which are usually modeled as Lipschitz functions. In [48], the gain adaptation allows one to account for the unknown Lipschitz constant. Resetting rules are proposed in [49]. The use of a time-varying gain is addressed in [50,51], where a Lyapunov functional is used for the purpose of the stability analysis of the estimation error instead of the classical quadratic Lyapunov function.

A new high-gain observer able to overtake some of the drawbacks of classical structures has been recently proposed in [52] for a class of nonlinear systems with one output and dimension $n \geq 3$. The cornerstone of this contribution consists in limiting the power of the observer gain to 2 regardless of the dimension of the system, thus improving the performance of the observer with respect to the measurement noise on the output. Although the new observer structure solves the problem of numerical implementation, the peaking phenomenon is still present. Along this route, two similar schemes, which follow the seminal idea presented in [52], have been recently proposed, in [53] and [54], to address the implementation issues and the peaking phenomenon. In [53], the author shows how to build a high-gain observer by interconnecting a cascade of reduced-order high-gain observers of dimension 1. A simpler scheme, without feedback interconnection terms, that cannot ensure asymptotic estimate, is presented in [54]. It is worth stressing, however, that even if the dimension of the observers is $n$, neither scheme improves the sensitivity properties with respect to standard high-gain observers.

The first objective of this thesis is to tackle the challenging performance issues that arise when implementing high-gain observers. We develop a new high-gain observer design method for nonlinear systems that has a lower gain compared to the standard high-gain observer in addition to reducing the sensitivity to noise measurement. The idea is to augment the dimension of the system such that the nonlinearity does not depend on the last components of the augmented system and explore the HG/LMI observer developed in [55] to decompose the nonlinearity of the system which allows reducing the Lipschitz constant directly and proportionally related to the high-gain tuning parameter.

The problem of observer design of nonlinear systems has been addressed from a different approach using moving horizon estimation techniques (MHE) [56–59]. The objective of these methods is to provide numerical solutions instead of analytical ones. These solutions are obtained from the resolution, in the sense of least squares, of a system of nonlinear equations using optimization algorithms. In the second part of the thesis, we study the problem of state estimation for quasi-LPV systems using a moving horizon estimator, this latter will be used to estimate both state and variable parameters of the system simultaneously, using the pessimistic and optimistic approaches.

## Thesis Organization

**Chapter 1** is essentially devoted to a presentation of some essential preliminaries necessary for the understanding of this thesis. A reminder of the notions of stability and observability for nonlinear systems is provided in the first section followed by a state-of-the-art on the different existing methods concerning the design of observers for nonlinear systems in a non-exhaustive way.

In **chapter 2**, the reader can find an overview of the theory of high-gain observers for nonlinear systems where the main features and the main drawbacks are highlighted. Then, we introduce a novel methodology for the design of high-gain observers by exploiting the HG/LMI technique recently developed in [55] and the system state augmentation approach that helps in overcoming (or improving) some of the main drawbacks of the standard high-gain observers. First, particular attention is given to the design of the high-gain observer using the state augmentation approach, where we highlight the benefit of augmenting the state of the system in achieving a desired (fast) state reconstruction without sacrificing the steady-state performance in the presence of measurement noise. The obtained results are then combined with HG/LMI technique which gives us more degrees of freedom in selecting the observer parameter, thus getting a good tradeoff between fast state reconstruction and measurement noise attenuation. We devote **chapter 3** to the problem of state estimation using the moving horizon estimation technique (MHE). We focus on estimation for nonlinear systems that can be written under the form of quasi-linear-parameter-varying systems (quasi-LPV) with bounded unknown parameters. Moving-horizon estimators are proposed to estimate the state of such systems according to two different formulations, i.e., "optimistic" and "pessimistic". In the former case, we perform estimation by minimizing the least-squares moving-horizon cost with respect to both state variables and parameters simultaneously. In the latter, we minimize such a cost with respect to the state variables after picking up the maximum with respect to the parameters. The analysis of the convergence of the MHE estimator will be discussed.

**Chapter 4** is dedicated to the validation of the obtained results on biological applications, namely, genetic regulatory network (GRN), SI epidemic model, and dorsal closure of Amnioserosa cells. First, a brief presentation of the models under study is given with the motivation to perform state estimation on these models. Simulation results are reported at the end of the chapter to demonstrate the applicability of the proposed high-gain observer and MHE.
This dissertation closes with an overall conclusion highlighting the various points raised in the four chapters and offering a set of perspectives for future work.

Chapter 1

# An overview on observer tools for nonlinear systems

*"Measure what is measurable and make it measurable what is not so."*

Galileo Galilei

## Contents

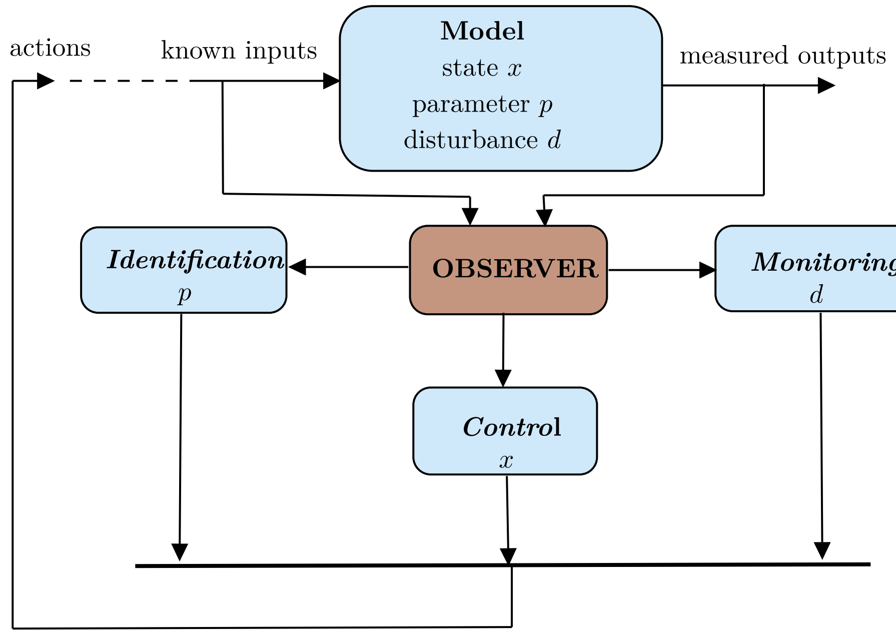



Figure 1.1: Observer as the heart of control systems

## 1.1 Intoduction

The problem of observer design naturally arises in a system approach, as soon as one needs some internal information from external (directly available) measurements. This need for internal information can be motivated by various purposes: modeling (*identification* ), monitoring (*fault detection*), or driving (*control*) the system. All those purposes are jointly required when aiming at keeping a system under control, as summarized by figure 1.1 hereafter. This makes the reconstruction -or observer-problem the heart of a general control problem.

The purpose of this chapter is to introduce the problem of observer design for nonlinear systems and presents some basic notions of observability and stability which will be needed throughout the thesis. In this framework, section 1.2 provides a set of definitions concerning stability (in the Lyapunov sense) and observability of nonlinear dynamic systems. Unlike linear systems, the observability of nonlinear systems is related to inputs and initial conditions, hence the fundamental role played by inputs in the study of the observability of nonlinear models is highlighted, and the concepts of universal inputs and uniformly observable systems are introduced in section 1.3. Section 1.4 provides state-of-the-art about different observer design techniques for nonlinear systems. We will see that there is no universal method for the synthesis of observers, and the possible approaches are either an approximation of linear algorithms or specific nonlinear algorithms. We present the main approaches developed in this field since the 1970s, a field which nevertheless remains very open, in particular, because of the multiplicity and diversity of nonlinear systems.

## 1.2 On the stability of dynamical systems: Lyapunov stability

In this section, we recall some fundamental concepts on the stability of continuous-time and discrete-time dynamical systems. By definition, the analysis of the stability of a dynamic system amounts to studying its behavior (its trajectory) when its initial state is close to a point of equilibrium [60, 61]. The main notions of stability are presented here, namely the asymptotic stability, exponential stability, in the sense of Lyapunov's theory.

### 1.2.1 Stability of continuous-time systems

Consider a dynamic system that satisfies:

$$\dot{x}(t) = f(x,t), \quad x(t_0) = x_0, \quad x(t) \in \mathbb{R}^n \tag{1.1}$$

We will assume that $f(x,t)$ satisfies the standard conditions for the existence and uniqueness of solutions. Such conditions are, for instance, that $f(x,t)$ is Lipschitz continuous with respect to $x$, uniformly in $t$, and piecewise continuous in $t$. A point $x_e$ is an *equilibrium point* of (1.1) if $f(x_e,t) = 0$, $\forall t \geq t_0$, $x(t_0)$ and $t_0$ are the initial state and initial time, respectively. We denote by $x(t,t_0,x_0)$ the solution of system (1.1) at $t \geq t_0$.
Throughout this thesis, only the stability of the estimation error is studied. For this reason, we assume that the nonlinear system (1.1) possesses a unique equilibrium point $x_e = 0$. This assumption leads to representing some definitions of the stability of system (1.1) at the origin.

**Definition 1.1: Stability**

The equilibrium point $x_e = 0$ (origin) of the system (1.1) is said to be *stable* (*in the sense of Lyapunov*) at $t = t_0$ if for any $\epsilon > 0$, there exists a positive scalar $\delta(\epsilon, t_0) > 0$ such that

$$\|x(t_0)\| < \delta(\epsilon, t_0) \Rightarrow \|x(t,t_0,x_0)\| < \epsilon, \ \forall t \geq t_0 \geq 0 \tag{1.2}$$

The system (1.1) is said to be unstable if it is not stable.

Lyapunov stability is a very mild requirement on equilibrium points. In particular, it does not require that trajectories starting close to the origin tend to the origin asymptotically. Also, stability is defined at a time instant $t_0$. *Uniform stability* is a concept that guarantees that the equilibrium point is not losing stability. We insist that for a uniformly stable equilibrium point $x_e$, $\delta$ in the Definition 1 not be a function of $t_0$, so that equation 1.2 may hold for all $t_0$ as defined in the following.

**Definition 1.2: Uniform stability**

The equilibrium point $x_e = 0$ of system (1.1)is said to be uniformly stable (in the sense of Lyapunov), if for any $\epsilon > 0$, there exists a positive scalar $\delta(\epsilon)$ such that:

$$\|x(t_0)\| < \delta(\epsilon) \Rightarrow \|x(t,t_0,x_0)\| < \epsilon, \ \forall t \geq t_0 \geq 0 \tag{1.3}$$

**Definition 1.3: Asymptotic stability**

The equilibrium point $x_e = 0$ of system 1.1 is *asymptotically stable* at $t = t_0$, if

1. $x_e = 0$ is stable, and

2. $x_e = 0$ is locally attractive, i.e., there exists a positive scalar $\delta(t_0)$ such that:

$$\|x(t_0)\| < \delta(t_0) \Rightarrow \lim_{t\to\infty} \|x(t, t_0, x_0)\| = 0, \ \forall t \geq t_0 \geq 0 \tag{1.4}$$

As in the previous definition, asymptotic stability is defined at $t_0$. Uniform asymptotic stability requires:

1. $x_e = 0$ is uniformly stable, and
2. $x_e = 0$ is uniformly locally attractive, i.e., there exists $\delta$ independent of $t_0$ for which equation (1.4) holds. Further, it is required that the convergence in (1.4) is uniform.

Definitions 1, 2, and 3 are local definitions; they describe the behavior of a system near an equilibrium point. We say an equilibrium point $x_e$ is *globally* stable if it is stable for all initial conditions $x_e \in \mathbb{R}^n$. Global stability is very desirable, but in many applications, it can be difficult to achieve. Notions of uniformity are only important for time-varying systems. Thus, for time-invariant systems, stability implies uniform stability and asymptotic stability implies uniform asymptotic stability.

**Definition 1.4: Exponential stability**

The equilibrium point $x_e = 0$ is an *exponentially stable* equilibrium point of of system (1.1) if there exist constants $m, \alpha > 0 \text{and} \epsilon > 0$ such that

$$\|x(t)\| \leq m \exp\left(-\alpha\left(t - t_0\right)\right) \tag{1.5}$$

for all $t \geq t_0$, $\|x(t_0)\| \leq \epsilon$.. The largest constant $\alpha$ which may be utilized in (1.5) is called the rate of convergence. The system is said to be *globally exponentially stable* if the bound in equation (1.5) holds for all $x_e \in \mathbb{R}^n$.

Exponential stability is a strong form of stability; in particular, it implies uniform, asymptotic stability. Exponential convergence is important in applications because it can be shown to be robust to perturbations and is essential for the consideration of more advanced control algorithms.

Finally, we say that an equilibrium point is *unstable* if it is not stable.

### 1.2.2 Stability of discrete-time systems

Consider the discrete-time system described by the following difference equation

$$x(k+1) = f(x(k), k), \quad x(k_0) = x_0 \tag{1.6}$$

where $x(k) \in \mathbb{R}^n$ is the state vector, $f(x(k), k) : \mathbb{R}^n \times \mathbb{R}^+ \to \mathbb{R}^n$ is a continuous vector function, $x(k_0)$ and $k_0$ are the initial state vector and the initial time, respectively. We denote by $x(k, k_0, x_0)$ the solution of the difference equation (D.2) at $k \geq k_0$.

The stability definitions for continuous-time systems (1.1) remain valid for discrete-time systems (1.6), except for the exponential stability which changes.

**Definition 1.5: Exponential stability**

The equilibrium point $x_e = 0$ of system (1.6) is said to be *locally exponentially stable*, if there exists positive scalars $m$, $\epsilon > 0$, and $0 < \rho < 1$ such that :

$$\|x(k)\| \leq m \|x(k_0)\| \rho^{(k-k_0)} \tag{1.7}$$

for all $k \geq k_0$, $\|x(k_0)\| \leq \epsilon$.. The largest constant $\alpha$ which may be utilized in (1.5) is called the rate of convergence. The system is said to be *globally exponentially stable* if the bound in equation (1.7) holds for all $x_e \in \mathbb{R}^n$.

### 1.2.3 Lyapunov direct method

The use of the previous definitions to check the stability of a system of the form (1.1) (resp. (1.6)) requires an explicit solution of the differential equation (1.1) (resp. the difference equation (1.6)). In most cases, it is not easy to compute the explicit solution of a nonlinear system or even impossible, which makes these definitions difficult to apply. as an alternative, Lyapunovâ€™s direct method (also called the second method of Lyapunov) allows us to determine the stability of a system without explicitly integrating the differential equation (1.1) (resp. (1.6)). The method is a generalization of the idea that if there is some "measure of energy" in a system, then we can study the rate of change of the energy of the system to ascertain stability. To make this precise, we need to define exactly what one means by a "measure of energy." Let $B_\epsilon$ be a ball of size around the origin, $B_\epsilon = x \in \mathbb{R}^n : \|x\| < \epsilon$ .

**Definition 1.6: Locally positive definite function (lpdf)**

A continuous function $V : \mathbb{R}^n \times \mathbb{R}_+ \to \mathbb{R}$ is a *locally positive definite function* if for some $\epsilon > 0$ and some continuous, strictly increasing function $\alpha : \mathbb{R}_+ \to \mathbb{R}$,

$$V(0,t) = 0 \quad \text{and} \quad V(x,t) \geq \alpha(\|x\|) \quad \forall x \in B_\epsilon, \forall t \geq 0.$$

A locally positive definite function is locally like an energy function. Functions that are globally like energy functions are called positive definite functions:

**Definition 1.7: Positive definite function (pdf)**

A continuous function $V : \mathbb{R}^n \times \mathbb{R}_+ \to \mathbb{R}$ is a *positive definite function* if it satisfies the conditions of Definition 6 and, additionally, $\alpha(p) \to \infty$ as $p \to \infty$.

To bound the energy function from above, we define decrescence as follows:

**Definition 1.8: Decrescent function**

A continuous function $V : \mathbb{R}^n \times \mathbb{R}_+ \to \mathbb{R}$ is decrescent if for some $\epsilon > 0$ and some continuous, strictly increasing function $\beta : \mathbb{R}_+ \to \mathbb{R}$,

$$V(x,t) \leq \beta(\|x\|) \quad \forall x \in B_\epsilon, \forall t \geq 0$$

**Definition 1.9: Lyapunov function**

A function $V(x,t)$ of class $\mathcal{C}^1$ is a local Lyapunov function (resp. global) for system (1.1), if it is proper positive definite and if there exists a subset $\mathcal{V}_0 \subset \mathbb{R}^n$ containing the origin, such that $\forall x \in \mathcal{V}_0$ (resp. $x \in \mathbb{R}^n$ ) :

$$\dot{V}(x,t) = \frac{\partial V(x,t)}{\partial t} + \frac{\partial V(x,t)}{\partial x} f(x(t),t) \leq 0.$$

If $\dot{V}(x,t) < 0$, then $V(x,t)$ is called a strict Lyapunov function for system (1.1).

Using these definitions, the following theorem allows us to determine the stability of a system by studying an appropriate Lyapunov function. Roughly, this theorem states that when $V(x,t)$ is a locally positive definite function and $\dot{V}(x,t) \leq 0$ then we can conclude the stability of the equilibrium point. In what follows, we denote by $\dot{V}$ the time derivative of $V$ along the trajectories of the system, $\dot{V}\Big|_{\dot{x}=f(x,t)}$.

**Theorem 1.1**

Let $V(x,t)$ be a non-negative function with derivative $\dot{V}$ along the trajectories of the system.

1. If $V(x,t)$ is locally positive definite and $\dot{V}(x,t) \leq 0$ locally in $x$ and for all $t$, then the origin of the system is locally stable (in the sense of Lyapunov).

2. If $V(x,t)$ is locally positive definite and decrescent, and $\dot{V}(x,t) \leq 0$ locally in $x$ and for all $t$, then the origin of the system is uniformly locally stable (in the sense of Lyapunov).

3. If $V(x,t)$ is locally positive definite and decrescent, and $-\dot{V}(x,t)$ is locally positive definite, then the origin of the system is uniformly locally asymptotically stable.

4. If $V(x,t)$ is positive definite and decrescent, and $-\dot{V}(x,t)$ is positive definite, then the origin of the system is globally uniformly asymptotically stable.

**Remark 1.1**

Theorem 1 gives sufficient conditions for the stability of the origin of a system. It does not, however, give a prescription for determining the Lyapunov function $V(x,t)$. Since the theorem only gives sufficient conditions, the search for a Lyapunov function establishing the stability of an equilibrium point could be arduous. However, it is a remarkable fact that the converse of Theorem 1 also exists: if an equilibrium point is stable, then there exists a function $V(x,t)$ satisfying the conditions of the theorem. However, the utility of this and other converse theorems is limited by the lack of a computable technique for generating

Lyapunov functions.

For the case of exponential stability:

**Definition 1.10: Exponential stability**

$x_e = 0$ is an *exponentially stable equilibrium point* of system (1.1) if and only if there exists an $\epsilon > 0$ and a function $V(x,t)$ which satisfies

$$\alpha_1\|x\|^2 \leq V(x,t) \leq \alpha_2\|x\|^2,$$
$$\dot{V}\Big|_{\dot{x}=f(x,t)} \leq -\alpha_3\|x\|^2,$$
$$\left\|\frac{\partial V}{\partial x}(x,t)\right\| \leq \alpha_4\|x\|$$

for some positive constants $\alpha_1, \alpha_2, \alpha_3, \alpha_4$, and $\|x\| \leq \epsilon$. The equilibrium point $x^* = 0$ is globally exponentially stable if the bounds in Theorem 4.5 hold for all $x$.

**Remark 1.2**

By choosing a quadratic Lyapunov function $V(x(t),t) = x^T(t)Px(t), P = P^T > 0$, then the origin of the linear system $\dot{x}(t) = Ax(t)$ is globally exponentially stable if $P$ is a solution for the matrix equation $A^TP + PA = -Q$, for any positive definite matrix $Q$.

Lyapunov's direct method can be applied to both continuous-time and discrete-time systems. The exponential stability of a discrete-time system is expressed as follows:

**Definition 1.11**

The origin of system (1.6) is locally exponentially stable if there exists a proper definite positive Lyapunov function $V(x_k,k) : \mathcal{B}_r \times \mathbb{R}^+ \rightarrow \mathbb{R}^+, V(0,k) = 0$, scalars $\alpha_1, \alpha_2$ et $0 < \alpha_3 < 1$ such that, $\forall x_0 \in \mathcal{B}_r$ et $\forall k \geq k_0 \geq 0$ :

1. $\alpha_1 \|x_k\|^2 \leq V(x,t) \leq \alpha_2 \|x_k\|^2$
2. The Lyapunov sequence $\{V(x_k,k)\}_{k=k_0,\ldots}$ is strictly decreasing, i.e :

$$\Delta V(x_k,k) = V(x_{k+1},k+1) - V(x_k,k) \leq -\alpha_3 V(x_k,k)$$

where

$$x_{k+1} = x(k+1,k_0,x_0) = f(x(k+1),k+1).$$

**Remark 1.3**

By choosing the Lyapunov quadratic function $V(x_k,k) = x_k^TPx_k, P = P^T > 0$, The origin of the discrete-time linear system $x_{k+1} = Ax_k$ is globally asymptotically stable, if and only if $P$ is a solution for the matrix equation $A^T PA - P = -Q$, for any positive definite matrix $Q$.

### 1.2.4 The indirect method of Lyapunov

The indirect method of Lyapunov uses the linearization of a system to determine the local stability of the original system. Consider the system defined by (1.1) with $f(0,t) = 0$ for all $t \geq 0$. Define

$$A(t) = \left.\frac{\partial f(x,t)}{\partial x}\right|_{x=0} \tag{1.8}$$

to be the Jacobian matrix of $f(x,t)$ with respect to $x$, evaluated at the origin. It follows that for each fixed $t$, the remainder

$$f_1(x,t) = f(x,t) - A(t)x \tag{1.9}$$

approaches zero as $x$ approaches zero. However, the remainder may not approach zero uniformly. For this to be true, we require the stronger condition that

$$\lim_{\|x\|\to 0} \sup_{t\geq 0} \frac{\|f_1(x,t)\|}{\|x\|} = 0. \tag{1.10}$$

If equation (1.10) holds, then the system

$$\dot{z} = A(t)z \tag{1.11}$$

is referred to as the (uniform) linearization of equation (1.1) about the origin. When linearization exists, its stability determines the local stability of the original nonlinear equation.

**Theorem 1.2: Stability by linearization**

Consider the system (1.1) and assume

$$\lim_{\|x\|\to 0} \sup_{t\geq 0} \frac{\|f_1(x,t)\|}{\|x\|} = 0.$$

Further, let $A(\cdot)$ defined in equation (1.8) be bounded. If 0 is a uniformly asymptotically stable equilibrium point of (1.11) then it is a locally uniformly asymptotically stable equilibrium point of (1.1).

This theorem proves that the global uniform asymptotic stability of the linearization implies the local uniform asymptotic stability of the original nonlinear system. The estimates provided by the proof of the theorem can be used to give a (conservative) bound on the domain of attraction of the origin. Systematic techniques for estimating the bounds on the regions of attraction of equilibrium points of nonlinear systems is an important area of research and involves searching for the "best" Lyapunov functions.

**Remark 1.4**

If the system (1.1) is time-invariant, then the indirect method says that if the eigenvalues of

$$A = \left.\frac{\partial f(x)}{\partial x}\right|_{x=0}$$

are in the open left half complex plane, then the origin is asymptotically stable.

### 1.2.5 Input-to-state stability

The concept of input-to-state stability was introduced by E. Sontag in his celebrated paper [62] as a test for the robustness of nonlinear systems to external perturbations, for a concise summary see [63]. In the framework of Input-to-state stability (ISS), error dynamics are regarded as a system whose input is the measurement noise which gives rise to disturbance-to-error stable (DES) observers. As a motivational example, consider the linear system described by

$$\dot{x} = Ax + Bu, \tag{1.12}$$

where $x(t) \in \mathbb{R}^n$ is the state, $u(t) \in \mathbb{R}^m$ the input and $A$ and $B$ constant matrices. The solution corresponding to the initial condition $x(0)$ can be written as

$$x(t) = \exp(At)x(0) + \int_0^t \exp(A(t-s))Bu(s)ds, \quad \forall t \geq 0 \tag{1.13}$$

and, if $A$ is Hurwitz, we can dominate the system state with

$$|x(t)| \leq c_1 \exp\left(-c_2 t\right)|x(0)| + c_3 \sup_{s\in[0,t]} |u(s)|, \quad \forall t \geq 0. \tag{1.14}$$

The positive constants $c_1, c_2$ and $c_3$ are such that

$$|\exp(At)| \leq c_1 \exp\left(-c_2 t\right), \quad c_3 = c_1 c_2^{-1}|B|$$

and, thus, they are independent of $x(0)$ and $u$. The first term in (1.14) quantifies the effect of the initial condition for short times, while the second term accounts for the input impact. These ideas can be generalized for nonlinear systems by using comparison functions.

**Definition 1.12**

A function $\gamma : \mathbb{R}^+ \to \mathbb{R}^+$is of class $\mathcal{K}$ if $\gamma(0) = 0$ and if it is strictly increasing and continuous. It is of class $\mathcal{K}_\infty$ if it is also unbounded. A function $\beta : \mathbb{R}^+ \times \mathbb{R}^+ \to \mathbb{R}^+$is of class $\mathcal{KL}$ if $\beta(r,t)$ is of class $\mathcal{K}$ for each $t \in \mathbb{R}^+$and if $\beta(r,t)$ decreases to zero as $t \to \infty$ for each $r \in \mathbb{R}^+$. We use the notation $\gamma \in \mathcal{K}$ or $\gamma \in \mathcal{K}_\infty$ and $\beta \in \mathcal{KL}$.

For the following, we consider a nonlinear system of the form

$$\dot{x} = f(x,u), \tag{1.15}$$

where $x(t) \in \mathbb{R}^n$ is the state, $u(t) \in \mathbb{R}^m$ the input and $f : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^n$ a continuously differentiable function such that $f(0,0) = 0$.

**Definition 1.13**

System (1.15) is said to be input-to-state practically stable (ISpS) if there exist functions $\beta \in \mathcal{KL}$ and $\gamma \in \mathcal{K}$ and a constant $c \geq 0$ such that for all inputs $u \in L_m^\infty$ and all initial conditions $x(0) \in \mathbb{R}^n$, the solution $x$ is defined on $\mathbb{R}^+$and it holds that

$$|x(t)| \leq \beta(|x(0)|,t) + \gamma\left(|u|_\infty\right) + c, \quad \forall t \geq 0. \tag{1.16}$$

The system is input-to-state stable (ISS) if the latter is satisfied with $c = 0$.

An alternative definition makes use of $\sup_{s\in[0,t]} \gamma(|u(s)|)$ instead of $\gamma(|u|_\infty)$, this is justified by the causality of the system. In such a case, the space of inputs can be defined by local boundedness $\left(L^\infty_{loc,m}\right)$.

**Remark 1.5**

We can recover a well-known stability property of system (1.15) if we fix $u = 0$. Indeed, the equilibrium point $x = 0$ is stable if for each $\epsilon > 0$ there exists $\delta > 0$ such that $|x(0)| < \delta \implies |x(t)| < \epsilon$, for all $t \in \mathbb{R}^+$. If additionally there exists $\delta > 0$ such that $|x(0)| < \delta \implies \lim_{t\to\infty} x(t) = 0$, we say the point is asymptotically stable. If for such a point the latter happens for all $x(0) \in \mathbb{R}^n$, then system (1.15) is called 0-globally asymptotically stable (0-GAS); which is equivalent to (1.16) with $u = 0$ and $c = 0$.

The concept of ISS was initially motivated by the problem of disturbances in state feedback stabilization. In fact, it is of interest to analyze the influence of a time-varying disturbance $d$ in the closed-loop system

$$\dot{x} = f(x, k(x) + d),$$

where $k$ stabilizes the system for $d = 0$. The author in [62] shows that, for control-affine systems, $k$ can be modified in order to achieve ISS when regarding $d$ as the system input. The elegant formulation of ISS has now become a standard tool in the literature in order to study the robustness of nonlinear systems with respect to inputs in a wide range of situations.
Showing directly the ISS of a system can be challenging. Therefore, the following definition and theorem provide a Lyapunov characterization of ISS.

**Definition 1.14**

A smooth function $V : \mathbb{R}^n \to \mathbb{R}^+$is called an ISpS-Lyapunov function for system (1.15) if there exist functions $\alpha_1, \alpha_2 \in \mathcal{K}_\infty$ and $\alpha_3, \chi \in \mathcal{K}$, and a constant $c_L \geq 0$ such that for all $x \in \mathbb{R}^n$ and all $u \in \mathbb{R}^m$ we have

$$\alpha_1(|x|) \leq V(x) \leq \alpha_2(|x|)$$

and $|x| \geq \chi(|u|) + c_L$ implies

$$\frac{\partial V}{\partial x}(x) f(x, u) \leq -\alpha_3(|x|).$$

The function is called ISS-Lyapunov if the latter holds with $c_L = 0$.

**Theorem 1.3**

System (2.32) is ISpS if and only if it has an ISpS-Lyapunov function. The same equivalence holds for ISS. As an example, the linear system (1.12) is ISS precisely when $A$ is Hurwitz. We can construct the ISS-Lyapunov function as $V(x) = x'Px$, where $A'P + PA < 0$. [64,65]

## 1.3 On the observability of dynamical systems

This section introduces the problem of observer design for nonlinear systems and presents some basic notions of observability that will be needed throughout the thesis. Our aim is not to provide an exhaustive study on nonlinear observability and observer design, but rather to situate our contribution and introduce the basic tools/notations needed in the rest of this thesis and give a brief overview on some existing techniques for designing observers for nonlinear systems.

### 1.3.1 Observation problem

We consider a general system of the form:

$$\dot{x} = f(x(t), u(t)), \quad y = h(x(t), u(t)) \tag{1.17}$$

where $x$ denotes the state vector, taking values in $X$ a connected manifold of dimension $n$, $u$ denotes the vector of known external inputs, taking values in some open subset $\mathcal{U}$ of $\mathbb{R}^m$, and $y$ denotes the vector of measured outputs taking values in some open subset $Y$ of $\mathbb{R}^p$.
Functions $f$ and $h$ will in general be assumed to be $\mathcal{C}^\infty$ with respect to their arguments, and input functions $u(.)$ to be locally essentially bounded and measurable functions in a set $\mathcal{U}$.
Given a model (1.17), the purpose of acting on the system, or monitoring it, will in general need to know $x(t)$, while in practice one has only access to $u$ and $y$. The observation problem can then be formulated as follows:

> *For any input $u$ in $\mathcal{U}$, any initial condition $x_0$ in $\mathcal{X}_0$, find an estimate $\hat{x}(t)$ of $x(t)$ based on the only knowledge of the input and output up to time $t$, namely $u_{[0,t]}$ and $y_{[0,t]}$, such that $\hat{x}(t)$ asymptotically approaches $x(t)$, at least when $\hat{x}(t)$ is defined on $[0, +\infty)$.*

Clearly, this problem makes sense when one cannot invert $h$ with respect to $x$ at any time.
In front of this, one can use the idea of explicit feedback in estimating $x(t)$. More precisely, if one knows the initial value $x(0)$, then an estimate of $x(t)$ can be obtained by simply integrating (1.17) from $x(0)$. Hence, if $x(0)$ is unknown, one can try to correct online the integration $\hat{x}(t)$ of (1.17) from some erroneous $\hat{x}(0)$ according to the measurable error $h(\hat{x}(t)) - y(t)$, namely to look for an estimate $\hat{x}$ of $x$ as the solution of a system:

$$\dot{\hat{x}}(t) = f(\hat{x}(t), u(t)) + k(t, h(\hat{x}(t)) - y(t)), \text{ with } k(t, 0) = 0 \tag{1.18}$$

Such an auxiliary system is what will be defined as an *observer,* and the above equation is the most common form of an observer for a system (1.17) (as in the case of linear systems [1,2]). A more rigorous mathematical definition is the following (a sketch is given in Figure 1.2).

**Definition 1.15**

An observer for (1.17) is given by an auxiliary system:

$$\dot{z}(t) = \Phi(z(t), u(t), y(t), t)$$
$$\hat{x}(t) = \Psi(z(t), u(t), y(t), t)$$

such that:

(i) $\hat{x}(0) = x(0) \Rightarrow \hat{x}(t) = x(t), \quad \forall t \geq 0$;

(ii) $\|\hat{x}(t) - x(t)\| \to 0$ as $t \to \infty$.

If (ii) holds for any $x(0), \hat{x}(0)$, the observer is global.
If (ii) holds with exponential convergence, the observer is exponential.
If (ii) holds with a convergence rate that can be tuned, the observer is tunable.

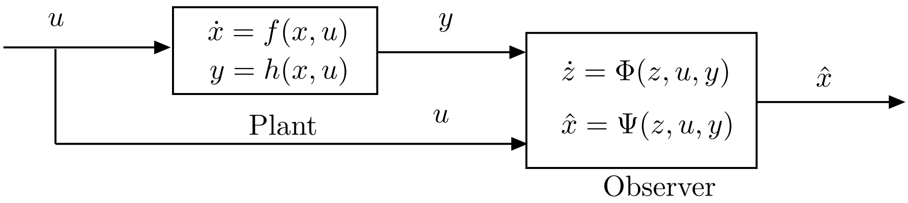



Figure 1.2: Observer: dynamical system estimating the state of a plant from the knowledge of its output and input only.

### 1.3.2 Observability for nonlinear systems

The role of an observer is to estimate the system state based on the knowledge of the input and output. This means that those signals somehow contain enough information to uniquely determine the system's whole state. This brings us to the notion of *observability*. In the case of nonlinear systems, the notion of observability is related to inputs and initial conditions. In this section, a more precise definition of observability will be given in the case of continuous-time systems of the form:

$$\begin{cases} \dot{x} = f(x,u), \\ y = h(x,u) \end{cases} \tag{1.19}$$

where $f : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^n$ and $h : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^p$
Observability is characterized by the fact that from an output measurement, one must be able to distinguish between various initial states, or equivalently, one cannot admit indistinguishable states (following [66]):

**Definition 1.16: Distinguishability and indistinguishability**

Two initial states $x_0$, $x_1 \in \mathcal{X}$ such that $x_0 \neq x_1$ are said to be distinguishable in $\mathcal{X}$ if $\exists t \geq 0$ and $\exists u : [0,t] \to \mathcal{U}$ an admissible input such that the trajectories of the outputs from $x_0$ and $x_1$, respectively, remain in $\mathcal{X}$ on the interval $[0,t]$, and satisfy $y(t, x_0, u(t)) \neq y(t, x_1, u(t))$. In this case, we say that the input $u$ distinguishes $x_0$ from $x_1$ in $\mathcal{X}$ . Conversely, two initial states $x_0$, $x_1 \in \mathcal{X}$ such that $x_0 \neq x_1$ are said to be indistinguishable if, $\forall t \geq$ and $\forall u : [0,t] \to \mathcal{U}$ for which the trajectories from $x_0$ and $x_1$ remain in $\mathcal{X}$, we have $y(t, x_0, u(t)) = y(t, x_1, u(t))$.

The notion of observability of a system at a single point [67,68] derives directly from the previous definition. By extension, it is possible to define the observability of a system at any point of $\mathcal{X}$.

**Definition 1.17: Observability**

The system (1.19) is observable at $x_0$ if for any other state $x_0 \neq x_1$, the two states $x_0$ and $x_1$ are distinguishable in $\mathcal{X}$. By extension, if this last property is true for any $x_0 \in \mathcal{X}$, then we say that the system is observable.

This last definition leads to the following theorem that can be found in [1].

**Theorem 1.4: [1]**

Let the continuous linear time-invariant system:

$$\begin{cases} \dot{x} = Ax(t) + Bu(t) \\ y = Cx(t) \end{cases} \tag{1.20}$$

with $A \in \mathbb{R}^{n \times n}$, $B \in \mathbb{R}^n$ and $C \in \mathbb{R}^n$. The system (1.20) is observable if and only if the observability matrix associated with this system is given by

$$\mathcal{O}_y = \begin{bmatrix} C \\ CA \\ CA^2 \\ \vdots \\ CA^{n-1} \end{bmatrix}$$

is of full rank. In this case, we say the pair $A, C$ is observable.

This result, widely used in the linear case, is in fact particularly restrictive in the nonlinear case. Indeed, the notion of observability as given above is a global result. However, in practice, we do not need to distinguish each trajectory on the set $\mathcal{X}$ and for any time interval $[t_0, t_0 + T[$. For this, we will recall in the following the notion of weak local observability, given by [66]. It is first necessary to define the notion of local observability [66].

**Definition 1.18: Local observability**

The system (1.19) is said to be locally observable at $x_0$ if, for any neighborhood $V_{x_0}$ of $x_0$, the set of indistinguishable states of $x_0$ in $V_{x_0}$ reduces to the singleton $x_0$. By extension, the system (1.19) is said to be locally observable if it is locally observable for all $x_0 \in \mathcal{X}$.

The notion of weak local observability given by [66] is then defined as follows.

**Definition 1.19: Local weak observability**

The system (1.19) is said to be locally weakly observable at $x_0$ if there exists an open neighborhood $V_{x_0}$ of $x_0$ such that for any open neighborhood $V'_{x_0} \subset V_{x_0}$ the set of indistinguishable states of $x_0$ in $V'_{x_0}$ reduces to the singleton $x_0$. By extension, the system (1.19) is said to be locally weakly observable if it is locally weakly observable for all $x_0 \in \mathcal{X}$.

Definition 19 means that a system is locally weakly observable if any state $x_0$ can be distinguished from its neighbors instantaneously. This notion is of more interest in practice and presents the advantage of admitting some *rank condition* characterization. Such a condition relies on the notion of *observation space* roughly corresponding to the space of all observable states.

**Definition 1.20: Observation space**

The observation space for a system (1.19) is defined as the smallest real vector space (denoted by $\mathcal{O}(h)$) of $\mathcal{C}^\infty$ functions containing the components of $h$ and closed under Lie derivation along $f_u(x) = f(x,u)$ for any constant $u \in \mathcal{U}$.

Note that the Lie derivative of $h^{(k)}(x)$ along the direction of the field $f_u$ at constant $u$ is given by

$$L_{f_u} h^{(k)}(x) = \frac{\partial h^{(k)}(x)}{\partial x} \cdot f_u(x) = \left(\nabla h^{(k)}(x)\right)^T \cdot f_u(x) = \frac{dy^{(k)}(t)}{dt} \tag{1.21}$$

and, more generally,

$$L^i_{f_u} h^{(k)}(x) = L_{f_u}\left(L^{i-1}_{f_u} h^{(k)}(x)\right) = \frac{d^i y^{(k)}(t)}{dt^i} \tag{1.22}$$

Let us now define the observability rank condition [66].

**Definition 1.21: Observability rank condition**

A system (1.19) is said to satisfy the observability rank condition at $x_0 \in \mathbb{R}^n$ if

$$\dim d\mathcal{O}(h)_{|x_0} = n \tag{1.23}$$

where, $d\mathcal{O}(h)_{|x_0} = d\psi(x_0), \psi \in \mathcal{O}(h)$.
Condition (1.23) is called the observability rank condition.
System (1.19) is observable if it satisfied the observability rank condition for any $x_0 \in \mathcal{X}$.
If system (1.19) satisfied the observability rank condition at $x_0$, then it is locally weakly observable in $x_0$.

In all the previous definitions, the impact of the input $u$ has not been taken into consideration. Indeed, in the case of nonlinear systems, the observability of a system strongly depends on the input applied to the system under consideration. The system (1.19) may be observable for some inputs and be not observable for other inputs. With this in mind, we will introduce the notion of universal input and that of *$\mathcal{U}$-uniform observability*, a key concept in the contributions presented in this manuscript [69].

**Definition 1.22: Universal input**

An admissible input $u : [0,T] \to \mathcal{U}$ is said to be universal for the system (1.19) on $[0,T]$ if, for any pair of distinct initial states $x_0, x_1$, there exists at least one instant of time $t \in [0,T]$ for which the resulting outputs of $x_0$ and $x_1$ are distinct, i.e. $y\left(t, x_0, u(t)\right) \neq y\left(t, x_1, u(t)\right)$.
A non-universal input is called a singular input.

In the case where all admissible inputs of $\mathcal{U}$ are universal, then any pair of initial states is distinguishable. This is called observability for any input, or $\mathcal{U}$-uniform observability [70].

**Definition 1.23: $\mathcal{U}$-uniform observability**

A system (1.19) for which all admissible inputs with values in $\mathcal{U}$ are universal is said to be $\mathcal{U}$-uniformly observable.

In the nonlinear case, there are several ways to define the notion of *observability*. For the concept of states indistinguishability, a very frequent definition has been established in [66]. Important results have been established in [67] and [71] for a special class of control-affine systems. For more details on the different types of definitions of the observability of nonlinear systems, we refer the reader to [66, 71–73].

**Remark 1.6**

The concept of observability mentioned above can be extended directly to the class of discrete-time systems. Different types of definitions have been discussed in [74].

## 1.4 Observers for nonlinear systems: a state of the art

The problem of state estimation for linear systems was completely solved in the 1960sâˆ'1970s by Kalman [1] in a stochastic approach and by Luenberger [2] in a deterministic framework. For the case of nonlinear systems, which concerns the majority of physical systems, the problem remains widely open giving rise to a wide range of estimation algorithms. The methods presented in the literature are either an extension of linear techniques which are based on a linearization of the model around an operating point or specific nonlinear algorithms [75, 76]. In this section, a fairly comprehensive, but by no means exhaustive, review of nonlinear state estimation methods will be given. All methods have their own positive and negative aspects, either as extensions of linear techniques or as novel nonlinear techniques.

### 1.4.1 Nonlinear transformation methods

This technique consists in transforming, using a change of coordinates, a nonlinear system into a linear system modulo an output injection. Once such a change of coordinates is obtained, the use of a Luenberger-type observer (corrected by output injection) can be used to estimate the state of the transformed system, and therefore the state of the original nonlinear system using the inverse change of coordinates. One of the first works in this direction is proposed in [4], where the autonomous system of the form

$$\dot{x} = f(x) \tag{1.24a}$$

$$y = h(x) \tag{1.24b}$$

is transformed by, a nonlinear change of coordinates $z = \Phi(x)$, to a linear system under the following observer canonical form

$$\dot{z} = A_c z + \lambda(y) \tag{1.25a}$$

$$y = C_c z \tag{1.25b}$$

where $A_c$ and $C_c$ are under the Brunovsky dual form, i.e.,

$$A_c = \begin{bmatrix} 0_{n-1} & I_{n-1} \\ 0 & 0_{n-1}^T \end{bmatrix}, \quad C_c = \begin{bmatrix} 1 & 0_{n-1}^T \end{bmatrix}.$$

The Luenberger observer corresponding to (1.24) is given by

$$\dot{\hat{z}} = A_c\hat{z} + \lambda(y) + K\left(y - C_c\hat{z}\right) \tag{1.26}$$

whose linear error dynamics $\varepsilon = z - \hat{z}$ is

$$\dot{\varepsilon} = \left(A_c - KC_c\right)\varepsilon. \tag{1.27}$$

The gain $K$ is obtained through a pole placement condition.
This method has been extended in [6] to the case of systems with multiple outputs and the nonlinear transformation has been generalized as follows

$$\begin{cases} z = \Phi(x) \\ v = \Psi(y) \end{cases} \tag{1.28}$$

where $v$ is the transformation of the output $y$ using the nonlinear change of coordinates $\Psi(.)$. The conditions under which such a transformation exists are established. However, three problems are related to this approach:

1. The class of systems for which such a transformation exists is very restrictive.
2. The procedure to find such a transformation is complicated.
3. In the case of multi-input systems (controlled systems), the transformed system contains all the input derivatives.

In [8], the system

$$\begin{aligned} \dot{x} &= f(x, u) \\ y &= h(x, u) \end{aligned} \tag{1.29}$$

is considered. In this case, the transformed system under the general canonical form is defined as

$$\dot{z} = A_c z + \lambda\left(y, u'\right) \tag{1.30a}$$

$$v = C_c z \tag{1.30b}$$

where $u' = \begin{bmatrix} u & \dot{u} & \dots & u^{(n)} \end{bmatrix}^T$. The nonlinear transformation used is

$$\begin{aligned} z &= \Phi\left(x, u'\right) \\ v &= \Psi\left(x, u'\right) \end{aligned} \tag{1.31}$$

By assuming that the derivatives of the input $u$ are available, the structure of the proposed observer is

$$\begin{aligned} \dot{\hat{z}} &= A_c\hat{z} + \lambda\left(y, u'\right) + K(v - \hat{v}) \\ \hat{v} &= C_c\hat{z} \end{aligned} \tag{1.32}$$

The dynamics of the estimation error are given by (1.27).
Other generalizations to multi-output systems have been proposed in [7,10] and [11]. A simplified algorithm to compute the suitable transformation, for the case of autonomous systems, has been designed in [9]. Necessary and sufficient conditions for the existence of the transformation for single-output systems have been given in [77]. These results have been generalized in [78] to systems with multiple outputs, and an algorithm to find the change of variables has been given. One reason why the class of systems that can be transformed into linear observable form is restricted is that the output must be linear as in (1.24b)and (1.30b). This condition is relaxed in [79] for the class of single-output autonomous systems. The idea is to transform the system (1.24), using the change of variables $z = \Phi(x)$, into

$$\dot{z} = Az + Ly \tag{1.33a}$$

$$y = \eta(z) \tag{1.33b}$$

where $\eta(z) = h(x)|_{x=\Phi^{-1}(z)}$. The obtained observer is

$$\dot{\hat{z}} = A\hat{z} + Ly, \tag{1.34}$$

and the error dynamics $\varepsilon = z - \hat{z}$ is

$$\dot{\varepsilon} = A\varepsilon. \tag{1.35}$$

The transformation $\Phi$ is chosen such that the matrix $A$ is obtained with the desired properties. In order to overcome the difficulty of obtaining the proper transformation, independently of previous works, a new approach has been presented in [5] for the class of autonomous and mono-output nonlinear systems. The necessary and sufficient conditions for the existence of the canonical form have been stated in [80]. Several extensions of this approach to the case of controlled and multi-output nonlinear systems are given in [81–83].

### 1.4.2 Extended observers

The observer based on the extended linearization method is another technique that exploits the useful tools available for linear systems. The gain of the observer is calculated from the linearized model around an operating point. This is for example the case of the extended Kalman filter and the extended Luenberger observer that we discuss in what follows.

#### Extended Kalman filter (EKF)

The extended Kalman filter is one of the most interesting and successful applications of the Kalman filter in the case of nonlinear systems. This extended filter consists in linearizing the equations of the standard Klaman filter by a first-order Taylor's formula.
The extended filter has been successfully applied to different types of nonlinear processes. Unfortunately, the proofs of stability and convergence established in the linear case, cannot be extended in a general way to the case of nonlinear systems. The analysis of the convergence of this estimator remains, at present, an open problem giving rise to a large number of papers and books [84–88].
Before introducing the extended Kalman filter, we need to introduce the standard Kalman filter for linear time-varying (LTV) systems.

- Continous time LTV systems: for the case of LTV systems in the form

$$\dot{x} = A(t)x + B(t)u + v_1(t) \tag{1.36a}$$

$$y = C(t)x + v_2(t) \tag{1.36b}$$

the standard Kalman filter is given by

$$\dot{\hat{x}} = A(t)\hat{x} + B(t)u + PC^T(t)R^{-1}(y - C(t)\hat{x}) \tag{1.37}$$

where $P$ is the symmetric positive definite solution of the following Riccati equation

$$\dot{P} = AP + PA^T + Q - PC^TR^{-1}CP. \tag{1.38}$$

- Discrete time LTV systems: for LTV systems in the form

$$x_{k+1} = A_k x_k + B_k u_k + v_k \tag{1.39a}$$

$$y_k = C_k x_k + w_k \tag{1.39b}$$

the standard Kalman filter is given by

$$\hat{x}_{k+1} = \hat{x}_{k+1/k} + K_{k+1}\left(y_{k+1} - C_{k+1}\hat{x}_{k+1/k}\right) \tag{1.40a}$$

$$P_{k+1} = \left(P_{k+1/k}^{-1} + C_{k+1}^T R_{k+1}^{-1} C_{k+1}\right)^{-1}; \tag{1.40b}$$

$$K_{k+1} = P_{k+1/k}C_{k+1}^T\left(C_{k+1}P_{k+1/k}C_{k+1}^T + R_{k+1}\right)^{-1} \tag{1.40c}$$

where

$$\hat{x}_{k+1/k} = A_k\hat{x}_k + B_k u_k \tag{1.41a}$$

$$P_{k+1/k} = A_k P_k A_k^T + Q_k \tag{1.41b}$$

$P_0 = \mu I_n > 0$. $\hat{x}_{k+1}$ and $\hat{x}_{k+1/k}$ are the estimation and prediction of the state $x_{k+1}$. The matrices $P_{k+1}$ and $P_{k+1/k}$ are the covariances of the estimation and prediction errors.. $Q_k$ and $R_{k+1}$ are weighting matrices depending on the stochastic variables $v_k$ and $w_k$.
The extended Kalman filter [89] is a direct extension of the standard Kalman filter by replacing the state and output matrices $A, C$ of system (1.36) or (1.39) by the jacobians of the system nonlinearities.
Consider the nonlinear system

$$\dot{x} = f(x,u) + v(t) \tag{1.42a}$$

$$y = h(x,u) + w(t) \tag{1.42b}$$

The EKF is described as follows

$$\dot{x} = f(\hat{x},u) + PH(\hat{x},u)R^{-1}(y - h(\hat{x},u)) \tag{1.43a}$$

$$\dot{P} = F(\hat{x},u)P + PF(\hat{x},u)^T + Q - PH(\hat{x},u)^TR^{-1}H(\hat{x},u)P \tag{1.43b}$$

where

$$F(\hat{x},u) = \frac{\partial f}{\partial x}(\hat{x},u)$$
$$H(\hat{x},u) = \frac{\partial h}{\partial x}(\hat{x},u).$$

In the case of discrete-time systems of the form

$$x_{k+1} = f(x_k, u_k) + G_k v_k \tag{1.44a}$$

$$y_k = h(x_k, u_k) + D_k w_k \tag{1.44b}$$

the EKF is given by

$$\hat{x}_{k+1} = \hat{x}_{k+1/k} + K_{k+1} e_{k+1} \tag{1.45}$$

where

$$P_{k+1} = (I_n - K_{k+1} H_{k+1}) P_{k+1/k} \tag{1.46a}$$

$$\hat{x}_{k+1/k} = f(\hat{x}_k, u_k) \tag{1.46b}$$

$$P_{k+1/k} = F_k P_k F_k^T + Q_k \tag{1.46c}$$

$$K_{k+1} = P_{k+1/k} H_{k+1}^T \left( H_{k+1} P_{k+1/k} H_{k+1}^T + R_{k+1} \right)^{-1} \tag{1.46d}$$

$$e_{k+1} = y_{k+1} - h\left( \hat{x}_{k+1/k}, u_{k+1} \right) \tag{1.46e}$$

$$F_k = F(\hat{x}_k, u_k) = \frac{\partial f}{\partial x}(\hat{x}_k, u_k) \tag{1.46f}$$

$$H_k = H(\hat{x}_k, u_k) = \frac{\partial h}{\partial x}(\hat{x}_k, u_k) \tag{1.46g}$$

where $P_0 = \mu I_n > 0$.
In common use, the matrices $Q_k$ and $R_{k+1}$ correspond to the process and measurement noise covariance matrices, respectively.

$$Q_k = G_k G_k^T, \quad R_{k+1} = D_{k+1} D_{k+1}^T.$$

**Extended Luenberger observer**

With reference to the extended Kalman filter algorithm, the extended Luenberger observer is proposed for nonlinear single-input single-output systems. The extended Luenberger observer is used, either on the original system with a constant gain or through a change of coordinates with a gain depending on the state to estimate. In the first case, a linearized model is needed, and the observer gain is calculated by pole placement. This type of observer can only be used when it is certain that the state will remain in the neighborhood of the equilibrium state. For this reason, this method is not widely because only local behavior can be guaranteed, i.e., the stability is guaranteed in a sufficiently small neighborhood of constant operating points. If disturbances and modeling errors are present, then performance and stability cannot be guaranteed. In the second case, as we mentioned previously, the methods based on a change of coordinates concern only a restricted class of nonlinear systems. Indeed, many approaches that use a change of coordinates require the integration of a set of nonlinear partial differential equations, which is often very tricky to achieve, therefore, only approximate solutions can be obtained. Some results and consequences for an extended Luenberger observer design are briefly summarized in [90].

### 1.4.3 Generalized Luenberger observer (GLO)

A new observer design technique has been proposed in [20,91–93]. The class of systems concerned by this new design is to add to the Luenberger observer, a second linear output feedback inside the nonlinear part of the system. This approach concerns systems described by the following equations:

$$\dot{x} = Ax + G\gamma(Hx) + \varrho(y, u) \tag{1.47a}$$

$$y = Cx. \tag{1.47b}$$

The proposed observer has the following structure

$$\dot{\hat{x}} = A\hat{x} + G\gamma(H\hat{x} + K(y - C\hat{x})) + \varrho(y, u) + L(y - C\hat{x}) \tag{1.48}$$

Convergence conditions of (1.48) have been established in [20]. This result concerns systems for which the nonlinear function $\gamma$ satisfies the following assumptions:

1. any component $\gamma_i$ is a scalar function with a scalar variable, i.e :

$$\gamma_i = \gamma_i\left(\sum_{j=1}^{j=n} H_{ij}x_j\right), i = 1, \ldots, r \tag{1.49}$$

3. all components of $\gamma$ are non-decreasing functions, i.e :

$$0 \leq \frac{\gamma_i(v) - \gamma_i(w)}{v - w}, \forall v \neq w \in \mathbb{R} \tag{1.50}$$

Using (1.47) and (1.48), the estimation error dynamics $\varepsilon = x - \hat{x}$ are given as

$$\dot{\varepsilon} = (A - LC)\varepsilon + G(\gamma(v) - \gamma(w)) \tag{1.51}$$

where

$$v = Hx \text{ et } w = H\hat{x} + K(y - C\hat{x}).$$

These convergence conditions are illustrated in the following theorem

**Theorem 1.5**

The estimation error (1.51) is exponentially stable at the origin if there exists a matrix $P = P^T > 0$, a constant $\nu > 0$ and a diagonal matrix $\Lambda > 0$ such that the inequality

$$\begin{bmatrix} (A - LC)^T P + P(A - LC) + \nu I_n & PG + (H - KC)^T \Lambda \\ G^T P + \Lambda(H - KC) & 0 \end{bmatrix} \leq 0 \tag{1.52}$$

is satisfied.

This technique has been extended in [19] and [93] to the case of monotonic multi-variable systems. Similar convergence conditions were obtained. New sufficient conditions for the synthesis of the gains $K$ and $L$ have been proposed in [92] for a class of systems whose nonlinearity is a scalar function with a scalar variable. This result is more general than the previous one since it takes into account the bounds of the term $\frac{\gamma(v)-\gamma(w)}{v-w}$ when they exist, i.e. when the nonlinearity satisfies the condition

$$0 \leq \frac{\gamma(v) - \gamma(w)}{v - w} \leq b, \forall v \neq w \in \mathbb{R} \tag{1.53}$$

By exploiting condition (1.53), the following theorem is derived.

**Theorem 1.6**

The observer (1.48) converges exponentially if there exists a matrix $P = P^T > 0$, a constant $\nu > 0$ and a diagonal matrix $\Lambda > 0$ such that the inequality

$$\begin{bmatrix} (A-LC)^T P + P(A-LC) + \nu I_n & PG + (H-KC)^T \\ G^T P + (H-KC) & -\frac{2}{b} \end{bmatrix} \leq 0 \tag{1.54}$$

is satisfied.

This last inequality is less restrictive than (1.52). In fact, in (1.52) it is necessary to have $PG + (H-KC)^T \Lambda = 0$ because of the presence of a zero on the diagonal which makes the inequality very restrictive. However, in (1.54), the zero on the diagonal is replaced by $-\frac{2}{b}$ which does not require $PG + (H-KC)^T$ to be null. Note that for $b = +\infty$, inequality (1.52) is obtained.

### 1.4.4 Triangular normal forms: high-gain designs

Triangular forms became of interest when [94] related their structure to uniformly observable systems, and when [95] introduced the phase-variable form for differentially observable systems. The celebrated high-gain observer proposed in [96] for phase variable forms and later in [26, 97] for triangular forms, have been extensively studied ever since.
The two papers [25, 98] are at the origin of these high-gain techniques which use Lyapunov's stability theory to adapt the techniques developed in the linear case. The method presented in [25] gives sufficient conditions for convergence of the estimated state towards the real state, for the class of nonlinear systems described by

$$\dot{x} = Ax + \phi(x,u) \tag{1.55a}$$

$$y = Cx \tag{1.55b}$$

where $x \in \mathbb{R}^n, u \in \mathbb{R}^m$ and $y \in \mathbb{R}^p$ represent the state vectors, inputs and outputs of the system, respectively. The pair $(A, C)$ is detectable and the nonlinearity, $\phi$ is Lipschitz with respect to $x$:

$$\|\phi(x,u) - \phi(\hat{x},u)\| \leq \gamma_\phi \|x - \hat{x}\|, \quad \forall x, \hat{x} \in \mathbb{R}^n \text{ et } \forall u \in \mathbb{R}^m \tag{1.56}$$

where $\gamma_\phi$ is the Lipschitz constant of the function $\phi$.
The *high-gain* observer has the following structure

$$\dot{\hat{x}} = A\hat{x} + \phi(\hat{x},u) + K(y - C\hat{x}) \tag{1.57}$$

The name "high-gain" comes from the structure of the observer: when the nonlinear function $f$ has a large Lipschitz constant, the difference between the real state and the estimated state increases. Therefore, the observer gain $L$ (1.57) must be large to compensate for this error amplification.
The dynamics of the estimation error $\varepsilon = x - \hat{x}$ is given by:

$$\dot{\varepsilon} = (A - LC)\varepsilon + \phi(x,u) - \phi(\hat{x},u) \tag{1.58}$$

The objective is to determine under which conditions the gain $L$ guarantees the stability of the estimation error $\varepsilon$ at zero.
Thau's method [25] provides a sufficient condition for the asymptotic stability of the estimation error (1.58). The result of this method is given by the following theorem.

**Theorem 1.7**

Consider system (1.55) and the observer (1.57). If the gain $L$ is chosen such that:

$$\gamma_\phi < \frac{\lambda_{\min}(Q)}{2\lambda_{\max}(P)} \tag{1.59}$$

where $\lambda_{\min}(S)$ and $\lambda_{\max}(S)$ are the minimum and maximum eigenvalues of the matrix $S$, respectively, the matrices $P = P^T > 0$ and $Q = Q^T > 0$ are solutions to the Lyapunov equation:

$$(A - KC)^T P + P(A - KC) + Q = 0 \tag{1.60}$$

then, the estimation error (1.58) is exponentially stable.

The proof of this theorem is based on the use of the standard Lyapunov function

$$V = V(\varepsilon) = \varepsilon^T P \varepsilon.$$

For more details on the proof of Theorem 7, see [25].
Thau's approach allows only to verify convergence of the observer (1.57), a posteriori. Indeed, the choice of the matrices $P$, $Q$ and $K$ which satisfy the inequality (1.59) is not direct. For example, placing the eigenvalues of $(A - LC)$ in the left half-plane does not imply that condition (1.59) is satisfied. There is no specific relationship between the eigenvalues of $(A - LC)$ and $\lambda_{\max}(P)$, this was proved in [16] by a simple numerical example.
Thau's method is not constructive, it gives no indication of the choice of a gain satisfying the condition (1.59). This is a verification technique, which guarantees the asymptotic convergence of the estimated state $\hat{x}$ towards the real state $x$ when the gain $L$ is already been chosen. The article [98] extended Thau's results in the deterministic framework. Among the so-called "high gain" techniques we can also find the work by [26, 99], and more recently [100], which propose to compensate for nonlinearity, at the level of the dynamics of the estimation error, by a sufficiently large gain (compared to the Lipschitz constant). An extension of these results to exponential observers is detailed in [30]. The work presented in [16] is at the origin of a series of articles describing constructive methods of the gain $L$ of the observer (1.57). Indeed, it proposes the following proposition:

**Proposition 1.1**

Consider system (1.55) and the observer (1.57). If there exists $\epsilon > 0$ such that the Riccati equation

$$AP + PA^T + P\left(\gamma_\phi^2 I_n - \frac{1}{\epsilon} C^T C\right) P + I_n + \epsilon I_n = 0 \tag{1.61}$$

admits a symmetric positive definite matrix $P$ as a solution, then the gain can be chosen as

$$L = \frac{1}{2\epsilon} P C^T \tag{1.62}$$

to ensure the asymptotic convergence of the observer (1.58).

However, this algorithm does not work for all observable pairs $(A, C)$ and unfortunately does not give information on the conditions to be satisfied by the matrix $(A - LC)$ in order to ensure

the stability of the estimation error. The placement of the eigenvalues of $(A - LC)$ in the left half-plane is certainly not enough.
In the work by [15], sufficient conditions on the matrix $(A - LC)$ have been established for the dynamical system (1.57) to be an observer of the system (1.55):

**Proposition 1.2**

System (1.57) is an observer for system (1.55) if the following conditions are satisfied:

(i) the pair $(A, C)$ is observable;

(ii) the gain $L$ is chosen such that $(A - LC)$ is stable and

$$\min_{\omega \in \mathbb{R}_+} \sigma_{\min}(A - LC - j\omega I) > k \tag{1.63}$$

The complete proof of this theorem is given in three steps in [15].
The work presented in [101] extended the previous results to reduced observers: it has been shown that the conditions of Proposition 2 also guarantee the existence of an asymptotically convergent reduced observer for the nonlinear system (1.55).
Other high-gain observer techniques have been developed in the literature, especially for the class of uniformly observable systems [26, 97, 102]. These methods use a change of variables to transform the system under consideration under the form of (1.55). The adaptive case is addressed in [28], which proposes an adaptive high-gain observer for nonlinear systems depending linearly on unknown parameters. The input-state stability theory is used in [103] to study its robustness to uncertainties. This type of observer has been applied to a class of biological systems and biotechnological processes in [26, 104, 105].

### 1.4.5 Variable structure observers

Variable structure observers constitute another family of observers. In all the previous methods, the studied system's dynamic model was assumed to be perfectly known. Here, it is a question of developing certain robustness with respect to parametric uncertainties. The method used to construct these observers is based on the theory of sliding modes [106, 107]. The class of systems studied is described by:

$$\begin{cases} \dot{x}(t) = Ax(t) + f(x(t), u(t)) \\ y(t) = Cx(t) \end{cases} \tag{1.64}$$

The function $f$ represents the nonlinearities and uncertainties of the system. The following assumptions are made about the system

(i) the pair $(C, A)$ is detectable, hence there exists a matrix $K$ such that the matrix $A - KC$ is stable;

(ii) the function $f$ is under the form

$$f(x(t), u(t)) = P^{-1}C^T h(x(t), u(t)) \tag{1.65}$$

where $P$ is a symmetric positive definite matrix, the solution to the Lyapunov equation ($P$ exists according to assumption (i))

$$(A - KC)^T P + P(A - KC) = -Q < 0 \tag{1.66}$$

the function $h$ is unknown but bounded

$$\|h(x(t),u(t))\| \leqslant \rho(u(t)) \tag{1.67}$$

We can underline that the nonlinearity does not appear in the structure of the observer proposed by [107]

$$\dot{\hat{x}}(t) = A\hat{x}(t) + K(y(t) - C\hat{x}(t)) + \kappa(\hat{x}(t),u(t),y(t)) \tag{1.68}$$

with

$$\kappa(\hat{x}(t),u(t),y(t)) = \begin{cases} \frac{P^{-1}C^T(y(t)-C\hat{x}(t))}{\|(y(t)-C\hat{x}(t))\|}\rho(u(t)) & \text{si } (x(t)-\hat{x}(t)) \neq 0 \\ 0 & \text{si } (x(t)-\hat{x}(t)) = 0 \end{cases} \tag{1.69}$$

The term $\kappa(\hat{x}(t),u(t),y(t))$ in (1.69) can be considered as a variable gain, which becomes infinite when the estimation error is small. It is shown in [75] that the observer (1.68) is an exponential observer of the system (1.64). It should be noted that the exact knowledge of the system is not necessary, it suffices to know an upper bound $\rho(u)$ on the nonlinearities or uncertainties. On the other hand, assumption (ii) imposes a structural constraint on $f$, which can be difficult to verify in the presence of model uncertainties. In [106], the authors propose a slightly different observer, which does not use this condition, but in return, global convergence is no longer guaranteed. The discontinuity of the function (1.69) is another drawback of this method: an oscillatory behavior can appear in the dynamics of the estimation error at high frequencies. To overcome this problem, the paper [108] proposes another choice for the function.

### 1.4.6 State estimation via online optimization

This class of estimation technique involves the formulation of state estimation as a minimization problem, wherein the state estimates are obtained by solving an online minimization problem [56–59, 109, 110]. The optimization is carried out over a horizon (into the past) using a series of continuously sampled measurements over time leading to a moving horizon state estimation. This estimation approach is in principle identical to Kalman filtering, however, the Kalman filter considers only one set of measurements at a time without taking into consideration the constraints on the system.
The following discrete-time linear system is considered,

$$x_{k+1} = Ax_k + Bu_k + Gw_k \tag{1.70a}$$

$$y_k = Cx_k + v_k \tag{1.70b}$$

for t = 0, 1, . . . , where $x_k \in \mathbb{R}^n$ is the state vector (the initial state $x_0$ is unknown) and $u_k \in \mathbb{R}^m$ is the control vector. The vector $w_k \in \mathbb{R}^n$ is an additive disturbance affecting the system dynamics. The state vector is observed through the measurement equation (1.70b), where $y_k \in \mathbb{R}^p$ is the observation vector and $v_k \in \mathbb{R}^p$ is a measurement noise vector.
In [111] is shown that the Kalman filter is the algebraic solution to the following unconstrained least-square optimization problem:

$$\min_{\hat{x}_0,\{\hat{w}\}_{k=0}^{T-1}} = \|\hat{x}_0 - \underline{x}_0\|^2_{P_0^{-1}} + \sum_{k=0}^{T-1} \|\hat{w}_k\|^2_{Q_k^{-1}} + \sum_{k=0}^{T} \|\hat{v}_k\|^2_{R_k^{-1}} \tag{1.71}$$

where

$$\begin{aligned} \hat{x}_{k+1} &= A\hat{x}_k + Bu_k + G\hat{w}_k \\ \hat{y}_k &= C\hat{x}_k + \hat{v}_k \end{aligned}$$

and $Q_k \succ 0, R_k \succ 0, P_0 \succ 0$ are positive definite matrices and $\underline{\mathrm{x}}_0$ is the mean of $\hat{x}_0$. This optimization problem now opens the possibility to add system knowledge in the form of constraints, then the optimization problem (1.71) becomes not equivalent to the Kalman filter anymore. If all the available past measurements are used for the estimation as in (1.71), the estimation problem grows unbounded with time which is referred to as the full information estimator [112]. In order to keep the estimation problem computationally tractable it is necessary to limit the processed data, for example by discarding the oldest measurement once a new one becomes available. This essentially slides a window over the data, leading to the moving horizon estimator (MHE). The data that is discarded can be accounted for by the so-called arrival cost so that the information is not lost. The MHE then considers only a limited amount of data so that the constrained optimization problem becomes:

$$\min_{\hat{x}_{T-N|T}, \hat{W}_T} \left\| \hat{x}_{T-N|T} - \underline{\mathrm{x}}_{T-N|T} \right\|^2_{P^{-1}_{T-N|T-1}} - \left\| Y^{T-1}_{T-N} - \mathcal{O}\hat{x}_{T-N|T} - \bar{c}bU^{T-2}_{T-N} \right\|^2_{\mathcal{W}^{-1}} + \sum_{k=T-N}^{T-1} \|\hat{w}_k\|^2_{Q_k^{-1}} + \sum_{k=T-N}^{T} \|\hat{v}_k\|^2_{R_k^{-1}} \quad (1.72)$$

such that,

$$\begin{aligned} \hat{x}_{k+1} &= A\hat{x}_k + Bu_k + G\hat{w}_k, \hat{y}_k = C\hat{x}_k + \hat{v}_k, \\ &\hat{x}_k \in \hat{\mathbb{X}} \triangleq \left\{ \hat{x}_k \in \mathbb{R}^n \mid \hat{D}_x \hat{x}_k \le \hat{d}_x \right\}, \\ &\hat{w}_k \in \hat{\mathbb{W}} \triangleq \left\{ \hat{w}_k \in \mathbb{R}^q \mid \hat{D}_w \hat{w}_k \le \hat{d}_w \right\} \\ &\hat{v}_k \in \hat{\mathbb{V}} \triangleq \left\{ \hat{v}_k \in \mathbb{R}^r \mid \hat{D}_v \hat{v}_k \le \hat{d}_v \right\} \end{aligned} \quad (1.73)$$

where $T$ is the current time, $Q_k \succ 0, R_k \succ 0, P_{T-N|T-1} \succ 0$ are the covariances of $w_k, v_k, x_{T-N}$ assumed to be symmetric, $N$ is the horizon length of the MHE, i.e. the amount of past data taken into account. $Y^T_{T-N} = \left[ y^T_{T-N}, \ldots, y^T_T \right]^T$ is a vector containing the past $N+1$ measurements, $U^{T-1}_{T-N} = \left[ u^T_{T-N}, \ldots, u^T_{T-1} \right]^T$ is a vector containing the past $N$ inputs. $x, w, v$ denote the variables of the system (1.70). $\hat{x}, \hat{w}, \hat{v}$ denote the estimated variables of system (1.73), and $\hat{x}^*_{T|T-N}$ and $\hat{W}^*_T = W^{T-1}_{T-N}{}^* = \{\hat{w}\}^{T-1}_{T|T-N}{}^*$ denote the optimizers of problem (1.72)-(1.73) where $\hat{W}_T = W^{T-1}_{T-N} = \{\hat{w}\}^{T-1}_{T|T-N}$ denotes the estimated noise sequence from time $T-N$ to time $T-1$. Finally, $\left\| \hat{x}_{T-N|T} - \underline{\mathrm{x}}_{T-N|T} \right\|^2_{P^{-1}_{T-N|T-1}} - \left\| Y^{T-1}_{T-N} - \mathcal{O}\hat{x}_{T-N|T} - \bar{c}bU^{T-2}_{T-N} \right\|^2_{\mathcal{W}^{-1}}$ is the arrival cost. For steady-state MHE$Q_k = Q, R_k = R$, and $P_{T-N|T-1} = P$ are time-invariant.
The current state of the system can be calculated from the initial state $x_{T|T-N}$ by forward programming using the system equations (1.70a) and (1.70b) if the deterministic input $U^{T-1}_{T-N}$ and the noise sequence $\{w\}^{T-1}_{T-N}$ are known. It is thus sufficient to estimate the initial state $\hat{x}^*_{T|T-N}$ and the noise $\hat{W}^*_T$.

**Remark 1.7**

In the case $T \le N$, the full information estimator is solved using the arrival cost:

$$\left\| \hat{x}_{T-N|T} - \underline{x}_{T-N|T} \right\|^2_{P_0^{-1}}.$$

The horizon "fills up" and no data is discarded [112].

This remark points out that wrongly posed constraints might lead to an infeasible optimization problem and that hard constraints on $\hat{v}_k$ could be problematic due to the possibility of outliers in the measurement. Any constraints posed in (1.73) should hence be chosen such that the real system does not violate them [112].
The MHE can only consider a limited amount of past data while remaining computationally tractable. The information in the discarded data should be preserved and this is achieved through the arrival cost which captures the older data through forward dynamic programming. It can hence be seen as equivalent to the cost to go into backward dynamic programming [112]. In the constrained case, the arrival cost cannot be calculated analytically. Caution needs to be taken because an ill-chosen arrival cost can lead to an unstable estimator, but in [112] it is also shown that the steady-state (constant $P_{T-N|T-1}$) constrained estimator is stable if $P_{T-N|T-1} \geq P_\infty$ where $P_\infty$ is the solution of the discrete algebraic Riccati equation:

$$P = A^T PA - A^T PB \left(B^T PB + R_{kal}\right)^{-1} B^T PA + Q_{kal} \tag{1.74}$$

The arrival cost needs to be updated at each time step. Two different update schemes are proposed in [113] and [112] for $\underline{\mathbf{x}}_{T-N|T}$ :

1. Filtered update scheme: use of the optimal estimate $N+1$ time steps in the past:

$$\underline{\mathrm{x}}_{T-N|T} = A\hat{x}^*_{T-N-1|T-N-1} + Bu_{T-N-1|T-N-1} + G\hat{w}^*_{T-N-1|T-N-1}$$

2. Smoothed update scheme: use of the optimal estimate from the last time step:

$$\underline{\mathrm{x}}_{T-N|T} = A\hat{x}^*_{T-N-1|T-1} + Bu_{T-N-1|T-1} + G\hat{w}^*_{T-N-1|T-1}$$

.

In [112], the author demonstrates that either update can give a better estimate than the full-information estimator if the estimation constraints (1.73) are not properly posed and the real system violates them but no general claim about robustness is made. The MHE with the smoothed update performs best but further research in this direction still remains open. The main disadvantage of the filtered update is that cycling effects can occur because the filtered update can be seen as $N$ independent parallel running filters [112,113]. The use of the smoothed update however avoids the cycling effect and hence this work only considers the smoothed update. The smoothed arrival cost is calculated as follows [114]:

$$\left\|\hat{x}_{T-N|T} - \underline{\mathrm{x}}_{T-N|T}\right\|^2_{P^{-1}_{T-N|T-1}} - \left\|Y^{T-1}_{T-N} - \mathcal{O}\hat{x}_{T-N|T} - \overline{cb}U^{T-2}_{T-N}\right\|^2_{\mathcal{W}^{-1}} \tag{1.75}$$

where for $i, j \leq N$

$$\begin{aligned} \mathcal{O} &= \begin{bmatrix} C^T & A^T C^T \dots A^{(N-1)^T} C^T \end{bmatrix}^T, \\ \mathcal{M}_{i,j} &= \begin{cases} 0 & \text{if } j \geq i, \\ CG & \text{if } j = i-1, \\ CA_{i-1}A_{i-2}\dots A_{j+1}G & \text{otherwise,} \end{cases} \\ \mathcal{W} &= \mathrm{diag}(R) + \mathcal{M}\,\mathrm{diag}(Q)\mathcal{M}^T \end{aligned} \tag{1.76}$$

and $P_{T-N|T-1}$ is calculated by the following backward Riccati equation [114]:

$$
\begin{aligned}
P_{k|T} &= P_{k|k} + P_{k|k}A^T P_{k+1|k}^{-1}\left(P_{k+1|T} - P_{k+1|k}\right)P_{k+1|k}^{-1}AP_{k|k},\\
P_{k|k} &= P_{k|k-1} - P_{k|k-1}C^T\left(R + CP_{k|k-1}C^T\right)^{-1}CP_{k|k-1}\\
P_{k|k-1} &= GQG^T + AP_{k-1|k-1}A^T
\end{aligned}
$$

Apart from the MHE formulation stated in equation (1.70a), a number of variations in the formulation can be found in the literature. In [115], the authors used the following objective function in which the estimates of the noise sequence $\hat{W}_T$ are not obtained: $J = \mu\left\|\hat{x}_{T-N|T} - \underline{\mathrm{x}}_{T-N|T}\right\|^2 + \sum_{k=T-N}^{T}\|y(T) - C\hat{x}(k \mid T)\|^2$. In this formulation, the scalar $\mu$ is used to guarantee stability. By reason of optimality, the formulation (1.70a) will give at least as good a result as the formulation in [115]. The authors in [116] introduced a nonzero mean $w_{m,k}$ for $\hat{w}_k$ using the filtered update of the arrival cost:

$$
\begin{aligned}
\min_{\hat{x}_{T-N|T},\hat{W}_{T-N|T}^{T-1}} &\left\|\hat{x}_{T-N|T} - \underline{\mathbf{x}}_{T-N|T}\right\|^2_{P^{-1}_{T-N|T-N}} + \\
&\sum_{k=T-N}^{T-1}\|\hat{w}_k - \mathbf{w_{m,k}}\|^2_{Q^{-1}} + \sum_{k=T-N}^{T}\|\hat{v}_k\|^2_{R^{-1}}
\end{aligned}
\tag{1.77}
$$

**Remark 1.8**

The unconstrained linear MHE is equivalent to the linear Kalman filter if the solution of the differential-algebraic Riccati equation (1.74) is used for $P_0$ in the arrival cost of the MHE and in the Kalman filter [112, 113].

## 1.5 Conclusion

The purpose of this chapter was to give an overview of fundamental concepts essential in observer design for nonlinear systems. The presentation follows a particular viewpoint on the problem and does not claim to be exhaustive. In particular, the most important notions of stability and observability (from this viewpoint) have been reviewed, and some observer techniques have been presented for discrete-time and continuous-time nonlinear systems. We have seen that the designs are driven by specific structures of systems and there is no universal method for the synthesis of nonlinear observers in the literature. The approaches developed to date are either an approximation of linear algorithms (linearization around an operating point) or specific nonlinear algorithms for certain classes of systems.

Chapter

2

# High-gain like observer design for a class of nonlinear systems

*"Accuracy of observation is the equivalent of accuracy of thinking."*

Wallace Stevens

## Contents



## 2.1 Introduction

High-gain observers play an important role in state estimation and output feedback control of nonlinear systems. After two seminal works appeared in 1992 [26, 27] the investigation of high-gain observers in nonlinear theory attracted the attention of many researchers. In the absence of measurement noise, this technique robustly estimates the derivatives of the output while achieving convergence of the estimation error that can be imposed arbitrarily fast by acting on a design parameter, appearing in the observer structure, typically known as the "high-gain parameter" [32]. Moreover, for a sufficiently high observer gain and a globally bounded controller, the high-gain observer is able to recover the system performance achieved with the state feedback control.

Although the effectiveness of such an observer has been nicely demonstrated in [117, 118], the standard high-gain observer faces a numerical challenge with high dimensional systems. Indeed, in the design of a standard high-gain observer, the observer gain is proportional to the powers of the tuning parameter denoted $\theta$ in this work, which is powered to the dimension of the observed state $n$. This creates a challenge in the numerical implementation when the state dimension is high or when the high-gain parameter has to be chosen largely to achieve fast estimation. Moreover, high-gain observers are known for having high sensitivity to high-frequency measurement noise, which makes state estimates practically unusable, especially for higher dimensional systems having nonlinearities with large Lipschitz constants.

Some of the earlier research performed in the spirit of high-gain observers can be found in the literature such as the so-called extended high-gain observer [42] which is composed of an extended high-gain observer (EHGO), for the estimation of the derivatives of the output, augmented with an extended Kalman filter (EKF) for the estimation of the states of the internal dynamics. Then, to account for the presence of disturbances acting on the system, several methods have been proposed based on gain adaptation methods [43], [44], [45], [46] and [47].

The selection of a high gain stems also from the need to account for the nonlinearities in the error dynamics, which are usually modeled as Lipschitz functions. In [48], the gain adaptation allows one to account for the unknown Lipschitz constant. Resetting rules are proposed in [49]. The use of a time-varying gain is addressed in [50], [51], where a Lyapunov functional is used for the purpose of the stability analysis of the estimation error instead of the classical quadratic Lyapunov function.

A new high-gain observer able to overtake some of the drawbacks of classical structures has been recently proposed in [52] for a class of nonlinear systems with one output and dimension $n \geq 3$. The cornerstone of this contribution consists in limiting the power of the observer gain to 2 regardless of the dimension of the system, thus improving the performance of the observer with respect to the measurement noise on the output. Although the new observer structure solves the problem of numerical implementation, the peaking phenomenon is still present. Along this route, two similar schemes, which follow the seminal idea presented in [52], have been recently proposed, in [53] and [54], to address the implementation issues and the peaking phenomenon. In [53], the author shows how to build a high-gain observer by interconnecting a cascade of reduced-order high-gain observers of dimension 1. A simpler scheme, without feedback interconnection terms, that cannot ensure an asymptotic estimate, is presented in [54]. It is worth stressing, however, that even if the dimension of the observers is $n$, neither scheme improves the sensitivity properties with respect to standard high-gain observers. Hence, the focus of this chapter is to analyze and address the issues that arise when implementing high-gain observers in both scenarios: noise-free and in the presence of measurement noise. In particular, we focus on the trade-off between fast-state reconstruction, minimizing the bound on the steady-state estimation error, and rejecting the high-frequency measurement noise.

First, we will present a new observer structure for triangular systems having Lipschitz nonlinearities. The proposed observer is based on system state augmentation which transforms the original system of dimension n into an augmented system of dimension $n + j_s$ which allows obtaining a new threshold on the observer parameter $\theta$ that guarantees the exponential convergence of the estimation error and reduces the value of the observer gain. Then, we combine the HG/LMI technique recently proposed in [55] with the system state augmentation approach to obtain a new enhanced high-gain observer. The key idea behind this proposed observer is based on transforming the original system of dimension $n$ into an augmented system of dimension $n + j_s$, then applying the HG/LMI technique to the resulting system. Such an observer has more degrees of freedom as compared with the standard high-gain observer, which can be regarded as a particular case of this improved high-gain observer because of a special choice of design parameters.

## 2.2 Highlights on high-gain observers

This section is devoted to presenting an overview of sufficient conditions for the existence of observability canonical forms for nonlinear systems. These observability forms can be seen as a special case of a feedback form. As a matter of fact, when the aforementioned sufficient conditions are verified, the nonlinear system is diffeomorphic to a system for which we know how to design an observer. For the sake of simplicity, we consider the class of single-input single-output nonlinear systems. The results presented herein cannot be extended to the multi-output multi-input case in a trivial way.

**Observabilility canonical forms**

Single input observability is the practical observability notion that can be used for state and parameter estimation. A system is single-input observable if there exists an input that distinguishes any different initial states (see chapter 1). Such inputs are called universal inputs. For analytic systems, the observability is equivalent to the single input observability [69]. For nonlinear systems, even if the system is single input observable, it may admit an input that makes it unobservable. However, for stationary linear systems, the single observability doesnâ€™t depend on the input and can be characterized using a Brunowsky canonical form [119]. The property that the single input observability does not depend on the input will be called uniform observability. As for stationary linear systems, canonical forms can be designed in order to characterize some class of uniformly observable nonlinear systems.
In the observation context, a natural extension of stationary linear systems consists in considering linear systems up to output injection:

$$\begin{cases} \dot{x} = Ax + \varphi((u, y)) \\ y = Cx \end{cases} \tag{2.1}$$

where the state $x = (x_1, \dots, x_n) \in \mathbb{R}^n$ evolves in compact subset $\mathcal{X}$ of $\mathbb{R}^n$, the input $u$ is any function assumed to be known evolving in compact subset $\mathcal{U}$ of $\mathbb{R}^m$ the known input and $y \in \mathbb{R}^p$ is the measured output.
The observability of $(C, A)$ is equivalent to the fact that system (2.1) is observable independently on the input.
An observer for systems (2.1) is a simple extension of the Luenberger observer:

$$\dot{\hat{x}} = A\hat{x} + \varphi((u, y)) + L(C\hat{x} - y) \tag{2.2}$$

where $L$ is any constant $n \times p$ constant matrix such that $A + LC$ is stable.
Based on this nice observability property and the fact that the observability is an intrinsic property (it does not depend on the system of coordinates), the problem of transforming a nonlinear system to systems of the form (2.1) by a change of coordinate has been initiated in [4] and extended to the multi-output case in [6, 7].

In the followings, we recall some of the results in [4].
Consider the single output nonlinear system:

$$\begin{cases} \dot{x} = f(x) \\ y = h(x) \end{cases} \tag{2.3}$$

where, $x(t) \in \mathbb{R}^n$, $u(t) \in \mathbb{R}^m$ and $y(t) \in \mathbb{R}^p$
In the single output case ($p = 1$), we will recall the necessary and sufficient condition that systems (2.3) must satisfy in order to be transformed into the canonical form:

$$\begin{cases} \dot{x} = Ax + \varphi(y) \\ y = Cx \end{cases} \tag{2.4}$$

where,

$$A = \begin{pmatrix} 0 & 0 & & 0 \\ 1 & 0 & & 0 \\ \vdots & \ddots & & \vdots \\ 0 & \dots & 1 & 0 \end{pmatrix} \text{ and } C = \begin{pmatrix} 0, & \dots, & 0, & 1 \end{pmatrix}$$

To do so, consider the family of vector fields $X_1, \dots, X_n$ defined by:

$$\begin{cases} L_{X_1}\left(L_f^k(h)\right) = 0, \text{ for } k = 0, \dots, n-1 \\ L_{X_1}\left(L_f^{n-1}(h)\right) = 1 \\ X_i = [X_{i-1}, f], \text{ for } i = 2, \dots, n \end{cases} \tag{2.5}$$

where $L_{X_1}$ denotes the Lie derivative along the vector field $X_1$ and [,] denotes the symbol of the Lie bracket operation.
Now, define the following transformation $\Phi = (\Phi_1, \dots, \Phi_n)$ by:
$L_{X_i}(\Phi_j)(x) = \delta_i^j$, where $\delta_i^j$ is the symbol of Kronecker.

**Theorem 2.1**

Assuming that the system (2.3) is observable in the rank sense at some $x_0 \in \mathbb{R}^n$. A necessary and sufficient condition for which $z = \Phi(x)$ becomes a local system of coordinates around $x_0$ in which system (2.3) becomes of the form (2.4) is that the vector fields $X_1, \dots, X_n$ commute. Namely, $[X_i, X_j] = 0$, for every, $i, j$.

For the proof (refer to [28]).
In [6, 7], the authors gave an extension of this result to the multi-output systems which can be transformed into the Brunowsky canonical form:

$$\begin{cases} \dot{x} = A_m x + \varphi(y) \\ y = C_m x \end{cases} \tag{2.6}$$

where,

$$A_m = \begin{pmatrix} 0 & 0 & & 0 \\ A_1 & 0 & & 0 \\ \vdots & \ddots & & \vdots \\ 0 & \dots & A_p & 0 \end{pmatrix}, A_i = \begin{pmatrix} 0 & 0 & & 0 \\ A_1 & 0 & & 0 \\ \vdots & \ddots & & \vdots \\ 0 & \dots & A_p & 0 \end{pmatrix}$$

is $n_k \times n_k$ matrix with $n_1 + \ldots + n_p$= n and $C_m = \begin{pmatrix} C_1 & 0 & 0 \\ \vdots & \ddots & \vdots \\ 0 & \ldots & C_p \end{pmatrix}$ with $C_p = \begin{pmatrix} 0, & \ldots, & 1 \end{pmatrix}$ a $n_k$ vector.

### 2.2.1 Canonical form and high-gain observer

Consider nonlinear systems of the form:

$$\begin{cases} \dot{x} = f(x, u) \\ y = h(x) \end{cases} \tag{2.7}$$

where, $x(t) \in \mathbb{R}^n$, $u(t) \in \mathbb{R}^m$ and $y(t) \in \mathbb{R}^p$ is the measured output.
System (2.7) is uniformly observable, if for every input $u \in L^\infty([0, T], \mathbb{R}^m)$, where $T > 0$ is fixed, $u$ is universal input. Namely, for every initial states $x_0, x_1$, the associated outputs $y(x_0, u, t), y(x_1, u, t)$ are not identically equal on $[0, T(, x_0, x_1, u)[$, where $T(, x_0, x_1, u) \leq T$ is largest time such that the trajectories $x(t)$ and $x_1(t)$ are will defined for every $t \in [0, T(, x_0, x_1, u)[$. Notice that if the linear part of systems (2.3) is observable then system (2.3) is uniformly observable. Moreover, an observer takes the form (2.2) observer exponentially converges whenever the unknown trajectory $x(t)$ is defined for all $t \geq 0$. A sufficient condition that guarantees the completeness of the system (ie. the trajectories are defined on the whole $\mathbb{R}^+$) is that $\varphi$ is a global Lipschitz function. Notice that completeness is necessary for the existence of an observer which converges as $t \to \infty$.

### 2.2.2 Observability canonical corm for uniformly observable systems

For the sake of simplicity, we consider the control affine nonlinear system:

$$\begin{cases} \dot{x} = f_0(x) + u_1 f_1 + \ldots + u_m f_m(x) \\ y = h(x) \end{cases} \tag{2.8}$$

where, the $f_i$ 's are assumed to be of class $\mathcal{C}^\infty$.
Given a function $\varphi$ from $\mathbb{R}^n$ into $\mathbb{R}$ of class $\mathcal{C}^n$, the the Lie derivatives of $\varphi$ along the vector $f_0$ are:
$L_{f_0}(\varphi) = \sum_{i=1}^n f_{0i} \frac{\partial \varphi}{\partial x_k}$. For $k = 1, \ldots, n$, $L_{f_0}^k(\varphi) = L_{f_0}\left(L_{f_0}^{k-1}(\varphi)\right)$, with $L_{f_0}^0(\varphi) = \varphi$.
Denote by $\Phi(x) = \begin{pmatrix} \Phi_1(x) \\ \vdots \\ \Phi_n(x) \end{pmatrix}$ the transformation defined by:
$\Phi_k(x) = L_{f_0}^{k-1}(h)(x)$, for $k = 1, \ldots, n$.
Then the following theorem is introduced.

**Theorem 2.2: [26]**

If system (2.8) is uniformly observable, then there exists an open dense subset $\mathcal{M}$ of $\mathbb{R}^n$ such that for every $x_0 \in \mathcal{M}$, there exists a neighborhood $V$, such that the map $\Phi$ becomes a diffeomorphism from $V$ into its range. Moreover, it transforms system (2.8) restricted to

$V$ into the following canonical form:

$$\begin{cases} \dot{z} = Az + \psi_0(z) + \sum_{i=1}^{m} \psi_i(z)u_i \\ y = Cz \end{cases} \tag{2.9}$$

$$A = \begin{pmatrix} 0 & 1 & 0 \\ \vdots & \ddots & \\ & & 1 \\ 0 & \dots & 0 \end{pmatrix}, \psi_0(z) = \begin{pmatrix} 0 \\ \vdots \\ \psi_n(z) \end{pmatrix}, \psi_0(z) = \begin{pmatrix} \psi_{k_1}(z_1) \\ \vdots \\ \psi_{k_j}(z_1, \dots, z_j) \\ \vdots \\ \psi_{k_n}(z_n) \end{pmatrix}, C = \begin{pmatrix} 1, & 0, & \dots, & 0 \end{pmatrix}.$$

Conversely, if a system (2.8) can be transformed into the above canonical form using any diffeomorphism, then the system is uniformly observable on the domain of definition of the diffeomorphism.

For the proof, we refer the reader to [26].

### High-gain observer design

In this section, we deal with the design of high-gain observers for the class of observer canonical form systems described as:

$$\begin{cases} \dot{z} = Az + \varphi(u, z) \\ y = Cz \\ z \in \mathbb{R}^n; \quad u \in \mathbb{R}^m \end{cases} \tag{2.10}$$

where, $A = \begin{pmatrix} 0 & 1 & 0 \\ \vdots & \ddots & \\ & & 1 \\ 0 & \dots & 0 \end{pmatrix}$ and $C = \begin{pmatrix} 1, & 0, & \dots, & 0 \end{pmatrix}$ and $\psi(z, u) = \begin{pmatrix} \psi_1(z_1, u) \\ \vdots \\ \psi_k(z_1, \dots, z_j, u) \\ \vdots \\ \psi_n(z, u) \end{pmatrix}$

In order to design the high-gain observer, the following assumption is first introduced.

**Assumption 2.1**

The nonlinear function $\varphi$ is a global Lipschitz function, i.e., for all bounded subset of $\mathbb{R}^m$; $\exists \gamma > 0$ such that $\forall z, z' \in \mathbb{R}^n$, we have $\|\varphi(z, u) - \varphi(z', u)\| \leq \|z - z'\|$, where $\|\|$ denotes the euclidian norm of $\mathbb{R}^n$.

**Remark 2.1**

This hypothesis guarantees the completeness of the system (for every admissible control, all trajectories of the system are defined in $\mathbb{R}^+$). If the concerned trajectories of the system lie into a bounded subset $\Omega$ of $\mathbb{R}^n$, then we can prolong the nonlinear term $\varphi$ to a global Lipschitz function $\tilde{\varphi}$ outside $\mathcal{B}$, so that trajectories of the new system coincide with those of the initial system.

Now, let $\theta > 0$ a parameter and set: $\Delta_\theta = \begin{pmatrix} \theta & 0 & & 0 \\ \vdots & \ddots & & \\ 0 & & \dots & \theta^n \end{pmatrix}$ and $L = \begin{pmatrix} l_1 \\ \vdots \\ l_n \end{pmatrix}$ such that $A + LC = \begin{pmatrix} l_1 & 1 & & 0 \\ \vdots & & \ddots & \\ l_{n-1} & 0 & & 1 \\ l_n & 0 \dots \dots & & 0 \end{pmatrix}$ becomes a Hurwitz matrix.

Then, the candidate observer has the following form:

$$\dot{\hat{z}} = A\hat{z} + \Delta_\theta L(C\hat{z} - y) \tag{2.11}$$

Then the following theorem is derived:

**Theorem 2.3**

Under assumption 1, system (2.11) forms an exponential observer for system (2.10). Let $\mathcal{U}$ be a compact subset of $\mathbb{R}^m$, then there exists a constant $\theta_0 > 0$ such that $\forall \theta > \theta_0; \exists \alpha > 0; \exists \beta > 0$ such that $\forall \hat{z}(0)$, we have $\|\hat{z}(t) - z(t)\| \leq \alpha e^{-\beta t} \|\hat{z}(0) - z(0)\|$, where $z(t)$ is the unknown trajectory to be estimated.

**Remark 2.2**

The high-gain observer (2.11) is characterized by the nice feature of being extremely easy to tune. The convergence of the observer can be arbitrarily chosen by selecting a large high-gain parameter $\theta$ in order to overcome the Lipschitz constant of the nonlinear function $\varphi$. More precisely, $\beta$ depends on the parameter $\theta$ and $\lim \beta(\theta) = +\infty$.

### The high-gain observer limitations

The high-gain observer design introduced in Theorem 3 highlights three main drawbacks of this approach:

1. **Implementation issues:** the high-gain observer (2.11) is characterized by having the gain of the output injection terms which is proportional to $\theta, \theta^2, \dots, \theta^n$. Furthermore, the minimum value of ‘ which guarantees asymptotic convergence of the observer, is proportional to the Lipschitz constant of the nonlinear function $\varphi$. As a consequence, if the high-gain parameter $\theta$ or the dimension n of the observed system is large, we need to implement in the observer a term ‘n which may be very harmful from a numerical point of view. If we want to avoid implementing powers of $\theta$ in the observer, we need some different strategies, such as a nonlinear change of coordinates, the use of non-linear functions, or dynamic extension.

2. **"peaking phenomenon":** convergence of the observer (2.11) has been stated in the proof of Theorem 3. In absence of measurement noise and of model uncertainties, the $\hat{z}$ dynamics can be bounded as
$$\|z(t) - \hat{z}(t)\| \leq \alpha e^{-\beta t} \|z(0) - \hat{z}(0)\|$$

During the transient, the decaying term $e^{-\beta t}$ is closed to one. As a consequence, the variable $\hat{z}$ shows a peak that is proportional to the error in the initial conditions and multiplied by a term $\alpha$, producing an estimate completely unreliable and which can be very large from a numerical point of view when $\theta$ are very large. The interaction of peaking with nonlinearities can induce finite escape time in output feedback scenarios. In particular, in the lack of global growth conditions, high-gain observers can destabilize the closed-loop system as the observer gain is driven sufficiently high. The peaking phenomenon has been extensively studied in literature [27, 37, 120] and different solutions have been proposed, based on re-scaling [27], projections [121], hybrid re-design [49] or time-varying gain approaches [122]. Finally, very recent publications [53] based on a nested-saturation design and [55] based on HG/LMI technique.

3. **Sensitivity to measurement noise:** one of the main features which questions the use of a high-gain observer in applications is its sensitivity to measurement noise making their use practically impossible in a really noisy environment. Indeed, the estimates may become completely unreliable, imposing some upper bound on the value of the high-gain parameter $\theta$ if estimation in presence of measurement noise is desired. This trade-off between the speed of the state estimation and the sensitivity to measurement noise is a well-known fact in the observer theory. In this respect, high-gain observers tuned to obtain fast estimation dynamics are necessarily very sensitive to high-frequency noise. Some strategies have been advanced in the literature to achieve fast convergence while reducing the impact of measurement noise at a steady state such as using a larger $\theta$ during the transient time and then decreasing it at steady state [118] and many other schemes [44–46, 55, 123].

In the next section, we propose a new analysis tool to overcome or at least to mitigate the aforementioned drawbacks.

## 2.3 Enhancing high-gain observer performances

This section addresses the challenging performance issues that arise when implementing high-gain observers in noise-free and in the presence of measurement noise. In particular, we focus on the trade-off between fast-state reconstruction, minimizing the bound on the steady-state estimation error, and rejecting the model uncertainty. Motivated by these considerations, we propose a new class of nonlinear high-gain observers, which substantially overtakes the drawbacks mentioned in the previous section. Our technique follows the standard high-gain methodology with the same state observer structure of dimension $n$. However, by exploiting the system state augmentation approach, we are able to decrease the tuning parameter (then implicitly, the gain power is decreased). This is achieved by augmenting the state of the system to the dimension $n + j_s$ where $j_s$ is a design parameter that can take values between 0 and $n$. Then, the HG/LMI technique proposed in [55] is combined with this state augmentation approach to avoid the peaking phenomenon, reduce the sensitivity to high-frequency measurement noise, and enhance the convergence rate if necessary. To get a good trade-off between all these criteria, the new observer offers the possibility to play with the values of $j_0$, $j_s$, and the tuning parameter $\theta$.

### 2.3.1 Basic ingredients of the observer construction

In this section, we introduce some background results on high-gain observers essential to the development of the proposed technique.

### System description

Before describing the high-gain observer form, it is important to lay the foundation for the types of systems that are considered in this body of work. Namely, We will consider the class of nonlinear systems in the triangular form described by the following set of equations:

$$\begin{cases} \dot{x} = \begin{pmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \vdots \\ \dot{x}_{n-1} \\ \dot{x}_n \end{pmatrix} = \begin{pmatrix} x_2 \\ x_3 \\ \vdots \\ x_n \\ f(x) \end{pmatrix} \\ y = x_1 \end{cases} \tag{2.12}$$

where $x \in \mathbb{R}^n$ is the system state, $x \in \mathbb{R}$ is the measured output, and $f : \mathbb{R}^n \to \mathbb{R}$ is a nonlinear function satisfying the Lipschitz property:

$$\Big| f(x_1 + \Delta_1, \ldots, x_n + \Delta_n) - f(x_1, \ldots, x_n) \Big| \leq k_f \sum_{j=1}^{n} |\Delta_j| \ . \tag{2.13}$$

Using a nonlinear transformation, system (2.12) can be rewritten under the form:

$$\begin{cases} \dot{x} = Ax + Bf(x) \\ y = Cx \end{cases} , \tag{2.14}$$

where the system matrices take the form

$$A = \begin{pmatrix} 0 & 1 & \ldots & \ldots & 0 \\ 0 & 0 & 1 & \ldots & 0 \\ \vdots & & & \ddots & \vdots \\ 0 & . & \ldots & 0 & 1 \\ 0 & 0 & \ldots & \ldots & 0 \end{pmatrix}, \ i.e., \ \ (A)_{i,j} = \begin{cases} 1 & \text{if} \quad j = i+1 \\ 0 & \text{if} \quad j \neq i+1 \end{cases} . \tag{2.15}$$

$$B = \begin{pmatrix} 0 & \ldots & 0 & 1 \end{pmatrix}^T, \ \ C = \begin{pmatrix} 1 & 0 & \ldots & 0 \end{pmatrix} \tag{2.16}$$

where $A \in \mathbb{R}^{n \times n}$, $B \in \mathbb{R}^{n \times 1}$ and $C \in \mathbb{R}^{1 \times n}$

It should be noticed that as demonstrated in [31], all uniformly observable systems can be transformed into system (2.14). Several real-world models are or can be transformed into the triangular form [31, 118]. The following references give more details about this family of systems and their practical importance [26, 28].

### Standard high-gain methodology

Here, we recall the basic standard high-gain observer as in [32]. For the class of nonlinear systems written in the triangular form as given in by equation (2.14), a candidate observer system is (as in [26]) just the "high-gain extended Luenberger observer":

$$\dot{\hat{x}} = A\hat{x} + Bf(\hat{x}) + L\Big(y - C\hat{x}\Big). \tag{2.17}$$

The dynamics of the estimation error $\tilde{x} = x - \hat{x}$ is then given by:

$$\dot{\tilde{x}} = (A - LC)\tilde{x} + B\big[f(x) - f(\hat{x})\big], \tag{2.18}$$

in which the observer gain $L$ is rewritten under the following form:

$$L := \mathrm{T}(\theta)K, \;\; \theta \geq 1. \tag{2.19}$$

where

$$\mathrm{T}(\theta) := \mathrm{diag}\Big(\theta, \ldots, \theta^n\Big) \text{ and } K \in \mathbb{R}^{n\times p}.$$

In addition, the high-gain methodology focuses on the transformed estimation error

$$\hat{\tilde{x}} := \mathrm{T}^{-1}(\theta)\tilde{x}, \tag{2.20}$$

where $\mathrm{T}^{-1}(\theta)$ is the inverse of $\mathrm{T}(\theta)$ given by

$$\mathrm{T}^{-1}(\theta) = \mathrm{diag}\Big(\frac{1}{\theta}, \ldots, \frac{1}{\theta^n}\Big).$$

It is well-known that the dynamics of the error $\hat{\tilde{x}}$ is given by

$$\dot{\hat{\tilde{x}}} = \theta(A - KC)\hat{\tilde{x}} + \frac{1}{\theta^n}B\Delta f, \tag{2.21}$$

with

$$\Delta f := f(x) - f(x - \mathrm{T}(\theta)\hat{\tilde{x}}).$$

From the Lipschitz condition (2.13) and the fact that $\theta \geq 1$, we can show as in [50] that there always exists a positive scalar constant $k_f$, independent of $\theta$, such that

$$\|\mathrm{T}^{-1}(\theta)B\Delta f\| \leq k_f\|\hat{\tilde{x}}\|. \tag{2.22}$$

then, the following theorem is derived,

**Theorem 2.4: [32]**

If there exist $P > 0$, $\lambda > 0$, $Y$ of appropriate dimensions, such that

$$A^T P + PA - C^T Y - Y^T C + \lambda I < 0, \tag{2.23}$$

then the observer converges exponentially to zero for

$$\theta > \max\{1, \frac{2k_f\lambda_{\max}(P)}{\lambda}\}, \tag{2.24}$$

and

$$K = P^{-1}Y^T$$

where $\lambda_{\max}(P)$ is the largest eigenvalue of the matrix $P$.

*Proof.* For more details about the proof of this theorem, we refer the reader to [32], [50], [51].
□

### 2.3.2 LPV-based approach

We consider a class of nonlinear systems without the linear state part which is not necessary for this design technique. For simplicity of presentation, we consider, without loss of generality, that the nonlinear function depends only on the system state, and the output is linear. The extension to nonlinear output is straightforward. Hence, the class of systems we treat in this section is given as

$$\begin{cases} \dot{x} = \Psi(x) \\ y = Cx \end{cases} \tag{2.25}$$

where the nonlinear function $\Psi : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ is assumed to be $\gamma_\Psi$-Lipschitz, i.e.:

$$\left\|\Psi(x) - \Psi(y)\right\| \leq \gamma_\Psi \left\|x - y\right\|, \quad \forall\; x, y \in \mathbb{R}^n \tag{2.26}$$

Hereafter, we introduce some definitions and preliminaries which will be of crucial use in the developed LPV-approach for Lipschitz and not necessarily differentiable systems.

**Definition 2.1**

Consider two vectors

$$X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in \mathbb{R}^n \quad \text{and} \quad Y = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} \in \mathbb{R}^n.$$

For all $i = 0, ..., n$, we define an auxiliary vector $X^{Y_i} \in \mathbb{R}^n$ corresponding to $X$ and $Y$ as follows:

$$\begin{cases} X^{Y_i} = \begin{pmatrix} y_1 \\ \vdots \\ y_i \\ x_{i+1} \\ \vdots \\ x_n \end{pmatrix} & \text{for } i = 1, ..., n \\ X^{Y_0} = X \end{cases} \tag{2.27}$$

Lemma 1 and 2 introduced in the following are crucial for the HG/LMI observer design.

**Lemma 2.1: [124]**

Consider a continuous function $\Psi : \mathbb{R}^n \longrightarrow \mathbb{R}$. Then, for all

$$X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in \mathbb{R}^n \quad \text{and} \quad Z = \begin{pmatrix} z_1 \\ \vdots \\ z_n \end{pmatrix} \in \mathbb{R}^n,$$

there exist functions $\psi_j : \mathbb{R}^n \times \mathbb{R}^n \longrightarrow \mathbb{R}, \quad j = 1, ..., n$ such that

$$\Psi(X) - \Psi(Z) = \sum_{j=1}^{j=n} \psi_j\Big(X^{Z_{j-1}}, X^{Z_j}\Big)\mathrm{e}_n^\top(j)\Big(X - Z\Big), \tag{2.28}$$

where $X^{Z_j} \in \mathbb{R}^n$ is an auxiliary vector corresponding to $X$ and $Z$ as follows:

$$\begin{cases} X^{Z_j} = \begin{pmatrix} z_1 \\ \vdots \\ z_j \\ x_{j+1} \\ \vdots \\ x_n \end{pmatrix} \text{ for } j = 1, ..., n \\ X^{Z_0} = X \end{cases} \tag{2.29}$$

where, $\mathrm{e}_n(j)$ is the $j^{th}$ vector of the canonical basis of $\mathbb{R}^n$.

*Proof.* The proof consists of rewriting $\Psi(X) - \Psi(Y)$ as

$$\Psi(X) - \Psi(Y) = \sum_{j=1}^{j=n} \Big[\Psi\Big(X^{Y_{j-1}}\Big) - \Psi\Big(X^{Y_j}\Big)\Big]$$

Now, defining the functions $\psi_j$ by

$$\psi_j\Big(X^{Y_{j-1}}, X^{Y_j}\Big) = \begin{cases} 0 & \text{if } \quad x_j = y_j \\ \frac{\Psi\left(X^{Y_{j-1}}\right) - \Psi\left(X^{Y_j}\right)}{x_j - y_j} & \text{if } \quad x_j \neq y_j \end{cases} \tag{2.30}$$

we can write

$$\begin{aligned} \Psi(X) - \Psi(Y) &= \sum_{j=1}^{j=n} \Big[\psi_j\Big(X^{Y_{j-1}}, X^{Y_j}\Big)\Big](x_j - y_j) \\ &= \sum_{j=1}^{j=n} \Big[\psi_j\Big(X^{Y_{j-1}}, X^{Y_j}\Big)e_n^T(j)\Big]\Big(X - Y\Big) \end{aligned} \tag{2.31}$$

□

**Lemma 2.2: [124]**

Consider a function $\Psi : \mathbb{R}^n \longrightarrow \mathbb{R}^n$, then, the two following items are equivalent:

- $\Psi$ is $\gamma_\Psi$-Lipschitz with respect to its argument, i.e.:

$$\Big\|\Psi(X) - \Psi(Z)\Big\| \leq \gamma_\Psi\Big\|X - Y\Big\|, \quad \forall\ X, Z \in \mathbb{R}^n; \tag{2.32}$$

- for all $i, j = 1, ..., n$, there exist functions

$$\psi_{ij} : \mathbb{R}^n \times \mathbb{R}^n \longrightarrow \mathbb{R}$$

and constants $\underline{\gamma}_{\psi_{ij}} \leq 0$, $\bar{\gamma}_{\psi_{ij}} \geq 0$, so that $\forall\ X, Z \in \mathbb{R}^n$,

$$\Psi(X) - \Psi(Z) = \sum_{i=1}^{i=n} \sum_{j=1}^{j=n} \psi_{ij} H_{ij}\Big(X - Z\Big), \tag{2.33}$$

and

$$-\gamma_\Psi \leq \underline{\gamma}_{\psi_{ij}} \leq \psi_{ij} \leq \bar{\gamma}_{\psi_{ij}} \leq \gamma_\Psi, \tag{2.34}$$

where

$$\psi_{ij} \triangleq \psi_{ij}\Big(X^{Z_{j-1}}, X^{Z_j}\Big) \text{ and } H_{ij} = \mathrm{e}_n(i)\mathrm{e}_n^\top(j).$$

*Proof.*

1. *Sufficiency:* we start by proving sufficiency. Assume that for all $i, j = 1, ..., n$, there exist functions

$$\psi_{ij}\ :\ \mathbb{R}^n \times \mathbb{R}^n \longrightarrow\ \mathbb{R}$$

and constants $\underline{\gamma}_{\psi_{ij}}$ and $\bar{\gamma}_{\psi_{ij}}$, so that (2.33) and (2.34) hold for all $X, Y \in \mathbb{R}^n$. Then, we have

$$\begin{aligned} \Big\|\Psi(X) - \Psi(Y)\Big\| &\leq \sum_{i=1}^{i=n} \sum_{j=1}^{j=n} |\psi_{ij}| \Big\|X - Y\Big\| \\ &\leq \left(\sum_{i=1}^{i=n} \sum_{j=1}^{j=n} \lambda_{ij}\right) \Big\|X - Y\Big\| \end{aligned} \tag{2.35}$$

where $\lambda_{ij} = \max\Big(|\underline{\gamma}_{\psi_{ij}}|, |\bar{\gamma}_{\psi_{ij}}|\Big)$. Hence, the function $\Psi$ is $\gamma_\Psi$-Lipschitz with

$$\gamma_\Psi \leq \sum_{i=1}^{i=n} \sum_{j=1}^{j=n} \max\Big(|\underline{\gamma}_{\psi_{ij}}|, |\bar{\gamma}_{\psi_{ij}}|\Big).$$

2. *Necessity:* We know that for all $X$, we can write

$$\Psi(X) = \begin{pmatrix} \Psi_1(X) \\ \vdots \\ \Psi_n(X) \end{pmatrix} = \sum_{i=1}^{i=n} e_n(i)\Psi_i(X)$$

Consequently, if $\Psi$ is $\gamma_\Psi$-Lipschitz, then we deduce that there are constants $0 < \gamma_{\Psi_i} \leq \gamma_\Psi$ for all $i = 1, ..., n$ so that each component $\Psi_i$ is $\gamma_{\Psi_i}$-Lipschitz. Indeed, we have

$$\begin{aligned} \Big\|\Psi(X) - \Psi(Y)\Big\|^2 &= \sum_{i=1}^{i=n} \Big|\Psi_i(X) - \Psi_i(Y)\Big|^2 \\ &\leq \gamma_\Psi^2 \Big\|X - Y\Big\|^2. \end{aligned} \tag{2.36}$$

Inequality (2.36) leads to

$$\Big|\Psi_i(X) - \Psi_i(Y)\Big| \leq \gamma_\Psi \Big\|X - Y\Big\|$$

which means that $\Psi_i$ is $\gamma_{\Psi_i}$-Lipschitz with $\gamma_{\Psi_i} \leq \gamma_\Psi$. From lemma 1, there are functions $\psi_{ij} \ : \ \mathbb{R}^n \times \mathbb{R}^n \longrightarrow \ \mathbb{R}$, $j = 1, ..., n$ so that

$$\Psi_i(X) - \Psi_i(Y) = \sum_{j=1}^{j=n} \psi_{ij}\Big(X^{Y_{j-1}}, X^{Y_j}\Big) e_n^T(j) \Big(X - Y\Big) \tag{2.37}$$

where $\psi_{ij}$ is given as in equation (2.30) of Lemma 1 by replacing $\Psi$ by $\Psi_i$. Since $\Psi_i$ is $\gamma_{\Psi_i}$-Lipschitz, then we have

$$\begin{aligned} \Big|\Psi_i(X^{Y_{j-1}}) - \Psi_i(X^{Y_j})\Big| &\leq \gamma_{\Psi_i} \Big\|X^{Y_{j-1}} - X^{Y_j}\Big\| \\ &= \gamma_{\Psi_i} \Big|x_j - y_j\Big| \end{aligned}$$

which means that

$$-\gamma_{\Psi_i} \leq \psi_{ij} \leq \gamma_{\Psi_i}.$$

This ends the proof.

□

Lemma 2 provides a best less conservative Lipschitz condition. Indeed, the reformulation (2.33)-(2.34) allows to treat the nonlinearity with the best precision and exploits all the interesting properties of the system's nonlinearity.

**LPV/LMI based observer**

Consider the following Luenberger observer:

$$\dot{\hat{x}} = \Psi(\hat{x}) + L\Big(y - C\hat{x}\Big) \tag{2.38}$$

The dynamic of the estimation error $e = x - \hat{x}$ is given by:

$$\dot{e} = \Big[\Psi(x) - \Psi(\hat{x})\Big] - LCe \tag{2.39}$$

Since $\Psi(.)$ is $\gamma_\Psi$-Lipschitz, then following Lemma 2 there are functions

$$\psi_{ij} \ : \ \mathbb{R}^n \times \mathbb{R}^n \longrightarrow \ \mathbb{R}$$

and constants $\underline{\gamma}_{\psi_{ij}}$ and $\bar{\gamma}_{\psi_{ij}}$, such that

$$\Psi(x) - \Psi(\hat{x}) = \Big[\sum_{i=1}^{i=n}\sum_{j=1}^{j=n} \psi_{ij} H_{ij}\Big] e \tag{2.40}$$

and

$$\underline{\gamma}_{\psi_{ij}} \leq \psi_{ij} \leq \bar{\gamma}_{\psi_{ij}} \tag{2.41}$$

where

$$\psi_{ij} \triangleq \psi_{ij}\Big(x_k^{\hat{x}_{j-1}}, x^{\hat{x}_j}\Big)$$

is defined as in (2.30),then replacing $\Psi$ by $\Psi_i$ (the $i^{\text{th}}$ component of $\Psi$).
For the sake of shortness, we use $\psi_{ij}$ instead of $\psi_{ij}\Big(x^{\hat{x}_{j-1}}, x^{\hat{x}_j}\Big)$ in what follows.

Now, define the matrices

$$\Theta = \Big(\psi_{ij}\Big)_{ij} \tag{2.42}$$

and

$$\mathcal{A}\Big(\Theta\Big) = \sum_{i=1}^{i=n}\sum_{j=1}^{j=n}\psi_{ij}H_{ij} \tag{2.43}$$

Consequently, the dynamics (2.39) can be rewritten as

$$\dot{e} = \Big[\mathcal{A}\Big(\Theta\Big) - LC\Big]e \tag{2.44}$$

According to (2.41), the matrix parameter $\Theta$ belongs to a bounded convex set $\mathcal{H}_n$ for which the set of vertices is defined by:

$$\mathcal{V}_{\mathcal{H}_n} = \Big\{\Phi \in \mathbb{R}^{n\times n} \ : \ \Phi_{ij} \in \Big\{\underline{\gamma}_{\psi_{ij}}, \bar{\gamma}_{\psi_{ij}}\Big\}\Big\}. \tag{2.45}$$

The following theorem is derived, it provides LMI conditions for the observer design of Lipschitz systems.

**Theorem 2.5: [124]**

The observer (2.38) is asymptotically convergent if there exist a positive definite matrix $\mathcal{P}$, a matrix $\mathcal{R}$ of appropriate dimension such that the following LMI conditions hold:

$$\mathcal{A}\Big(\Phi\Big)^T\mathcal{P} + \mathcal{P}\mathcal{A}\Big(\Phi\Big) - C^T\mathcal{R} - \mathcal{R}^TC < 0, \ \forall \ \Phi \in \mathcal{V}_{\mathcal{H}_n} \tag{2.46}$$

Hence, the observer gain is given by

$$L = \mathcal{P}^{-1}\mathcal{R}^T.$$

For the proof (refer to [124]).

**Remark 2.3**

The LPV/LMI-based approach is the best LMI technique which avoids high-gain, however, from the complexity point of view it is less interesting. Indeed, to synthesize the observer gain, the LPV/LMI-based approach typically needs to solve $2^{n^2}$ LMIs. In addition, this technique, as is the case for all LMI techniques, contrary to the high-gain method, provides sufficient LMI conditions from which we cannot guarantee the existence of a stable observer before solving the LMIs. On the other hand, the high-gain methodology guarantees convergence at the cost of a larger gain even for small values of the Lipschitz constants.

### 2.3.3 HG/LMI approach

To improve the design strategy based on the LPV approach, a combination of the high-gain methodology and the LPV-based technique is given in this section. The advantages of each method are exploited to get an improved observer design method. This latter is called "HG/LMI observer which proved a smaller gain in addition to a reduced number of LMIs conditions to be satisfied.

### Motivating example

The fact that $k_f$ in inequality (2.13) is independent of $\theta$ is not necessarily an advantage. Indeed, this depends on how $\theta$ would be involved in $k_f$. Also, the fact that $k_f$ is independent of $\theta$ does not come only from the condition $\theta \geq 1$, but essentially from the presence of the last component of $x$ in $f$. Because of this last component, the parameter $\theta$ vanishes from the term $\frac{1}{\theta^n}\Delta f$ for $\theta \geq 1$. In this brief and simple example, we present the motivation and the key idea of the HG/LMI technique. Consider a simple three-dimensional system.
If we take a nonlinear function

$$f(x) = \gamma_f \sin(x_3),$$

then, from (2.13) we get

$$\frac{1}{\theta^3}\|\Delta f\| \leq \frac{\gamma_f}{\theta^3} \times \mid \theta^3 \hat{\tilde{x}}_3 \mid = \gamma_f \mid \hat{\tilde{x}}_3 \mid \leq k_f \|\hat{\tilde{x}}\|,$$

where $k_f = \gamma_f$ in this case. However, if we take

$$f(x) = \gamma_f \sin(x_2),$$

then we get

$$\frac{1}{\theta^3}\|\Delta f\| \leq \frac{\gamma_f}{\theta^3} \times \mid \theta^2 \hat{\tilde{x}}_2 \mid = \frac{\gamma_f}{\theta} \mid \hat{\tilde{x}}_2 \mid \leq \frac{k_f}{\theta} \|\hat{\tilde{x}}\|.$$

Hence, by replacing in (2.24) $k_f$ by $\frac{k_f}{\theta}$, $\theta_0$ will be reduced to $\sqrt{\theta_0}$, which will reduce significantly the values of the observer gain.

### HG/LMI based observer

This observer design technique follows the standard high-gain methodology with the same state observer structure of dimension $n$. However, by exploiting the LPV/LMI presented in the previous section, the values of the tuning parameter and observer gain are decreased. Indeed, by introducing a "compromise index" $j_0$, with $0 \leq j_0 \leq n$, the power of the proposed high-gain is limited to $j_0$ with $2^{j_0}$ LMIs to solve instead of one single LMI as in standard high-gain observer.
We consider the standard high-gain observer structure:

$$\dot{\hat{x}} = A\hat{x} + f(\hat{x}) + L\Big(y - C\hat{x}\Big), \tag{2.47}$$

where $L$ is defined as in (2.19).
The dynamics of the transformed error $\hat{\tilde{x}}$, defined in (2.21), is then given by:

$$\dot{\hat{\tilde{x}}} = \theta(A - KC)\hat{\tilde{x}} + \mathrm{T}^{-1}(\theta)\Delta f \tag{2.48}$$

with

$$\Delta f := f(x) - f(x - \mathrm{T}(\theta)\hat{\tilde{x}}).$$

Each nonlinear component $f_i$ can be written under the form

$$\Delta f_i = \sum_{j=1}^{i-j_i} \theta^j \psi_{ij} \hat{\tilde{x}}_j + \sum_{j=1}^{j_i} \theta^{k_i(j)} \psi_{ik_i(j)} \hat{\tilde{x}}_{k_i(j)}, \tag{2.49}$$

where

$$k_i(j) = i - (j_0 - j),$$

$$0 \leq j_0 \leq i.$$

It follows that $\Delta f$ is written as

$$\Delta f = \underbrace{\overbrace{\sum_{i=1}^{n}\sum_{j=1}^{i-j_0} \theta^j \psi_{ij} \mathrm{e}_n(i) \hat{\tilde{x}}_j}^{\text{for HG}}}_{\Delta f_1} + \overbrace{\sum_{i=1}^{n}\sum_{j=1}^{j_0} \theta^{k_i(j)} \psi_{ik_i(j)} \mathrm{e}_n(i) \hat{\tilde{x}}_{k_i(j)}}^{\text{for LPV/LMI}}. \tag{2.50}$$

Hence, the error dynamics (2.21) is rewritten as follows:

$$\dot{\hat{\tilde{x}}} = \theta\Big(\mathcal{A}(\Psi^\theta) - KC\Big)\hat{\tilde{x}} + \mathrm{T}^{-1}(\theta)\Delta f_1, \tag{2.51}$$

where

$$\mathcal{A}(\Psi^\theta) = A + B\sum_{i=1}^{n}\sum_{j=1}^{j_i} \psi^\theta_{ij} \mathrm{e}_n(i)\mathrm{e}_n^\top(k_i(j)), \tag{2.52}$$

$$\Psi^\theta = \begin{pmatrix} \psi^\theta_{11} \\ \vdots \\ \psi^\theta_{1j_1} \\ \psi^\theta_{21} \\ \vdots \\ \psi^\theta_{2j_2} \\ \vdots \\ \psi^\theta_{nj_n} \end{pmatrix} \in \mathbb{R}^{\sum_{i=1}^{n} j_i}, \tag{2.53}$$

$$\psi^\theta_{ij} = \frac{\psi_{ik_i(j)}}{\theta^{1+(j_i-j)}}. \tag{2.54}$$

Now define the convex bounded set

$$\mathcal{H}^\sigma_{j_{\min}} = \left\{ \Phi \in \mathbb{R}^{\sum_{i=1}^{n} j_i} \;:\; \frac{\underline{\gamma}_{\gamma_{ik_i(j)}}}{\sigma^{1+(j_i-j)}} \leq \Phi_{ij} \leq \frac{\bar{\gamma}_{\gamma_{ik_i(j)}}}{\sigma^{1+(j_i-j)}} \right\} \tag{2.55}$$

for which the set of vertices is defined by

$$\mathcal{V}_{\mathcal{H}^\sigma_{j_{\min}}} = \left\{ \Phi \in \mathbb{R}^{\sum_{i=1}^{n} j_i} \;:\; \Phi_{ij} \in \left\{ \frac{\underline{\gamma}_{\gamma_{ik_i(j)}}}{\sigma^{1+(j_i-j)}}, \frac{\bar{\gamma}_{\gamma_{ik_i(j)}}}{\sigma^{1+(j_i-j)}} \right\} \right\}, \tag{2.56}$$

where $\underline{\gamma}_{\gamma_{ik_i(j)}} \leq 0$ and $\bar{\gamma}_{\gamma_{ik_i(j)}} \geq 0$ are respectively, the lower and upper bounds of the bounded parameter $\psi_{ik_i(j)}$. Since $\bar{\gamma}_{\gamma_{k(j)}} \geq 0$ and $\underline{\gamma}_{\gamma_{k(j)}} \leq 0$, then it is obvious that for two positive scalars $\sigma_1, \sigma_2$, we have the following implication

$$\sigma_1 < \sigma_2 \implies \mathcal{H}^{\sigma_1}_{j_0} \supset \mathcal{H}^{\sigma_2}_{j_0}. \tag{2.57}$$

Moreover,

$$\lim_{\sigma\to+\infty} \Big(\mathcal{H}^\sigma_{j_0}\Big) = \Big\{0_{\mathbb{R}^{j_0}}\Big\}. \tag{2.58}$$

On the other hand, we can show that there exists a positive real number $k_{j_0} \leq k_f$ such that $\Delta f_1$ satisfies

$$\|\mathrm{T}^{-1}(\theta)B\Delta f_1\| \leq \frac{k_{j_0}}{\theta^{j_0}}\|\hat{e}\|. \tag{2.59}$$

Hence, the following theorem is obtained.

**Theorem 2.6: [55]**

If there exist $P > 0$, $\lambda > 0$, $Y$, and $\sigma > 0$ such that

$$\mathcal{A}(\Psi^\sigma)^T P + P\mathcal{A}(\Psi^\sigma) - C^T Y - Y^T C + \lambda I < 0, \forall \Psi^\sigma \in \mathcal{V}_{\mathcal{H}^\sigma_{j_0}}, \tag{2.60}$$

$$\theta > \theta_{j_0}^{\frac{1}{1+j_0}} = \frac{2k_{j_0}\lambda_{\max}(P)}{\lambda}, \tag{2.61}$$

then the estimation error $\tilde{x}$ is asymptotically stable with

$$L = \mathrm{T}(\theta)\overbrace{P^{-1}Y^T}^{K},\ \theta \geq \max\left(\sigma, \theta_{j_0}^{\frac{1}{1+j_0}}\right).$$

*Proof.* For the proof, we refer the reader to [55]. □

### 2.3.4 New solution using system state augmentation approach

This section is devoted to the main result of this paper. The motivation of this work is inspired by the HG/LMI design presented in the previous section. We will show that by augmenting the state of the system, we can reduce the value of the tuning parameter and the power of the observer gain.

#### Motivation

The motivation for developing the new solution comes from work in [55]. Indeed, as demonstrated in [55], if the nonlinear function $f(.)$ satisfies the condition

$$\frac{\partial f}{\partial x_j}(x) \equiv 0, \forall\ j > n - j_s \tag{2.62}$$

for a given $j_s \geq 0$, then the Lipschitz inequality (2.22) becomes

$$\|\mathrm{T}^{-1}(\theta)B\Delta f\| \leq \frac{k_f}{\theta^{j_s}}\|\hat{\tilde{x}}\|. \tag{2.63}$$

It follows that the high-gain inequality (2.24) becomes

$$\theta > \left(\frac{2k_f\lambda_{\max}(P)}{\lambda}\right)^{\frac{1}{1+j_s}} = \theta_0^{\frac{1}{1+j_s}}. \tag{2.64}$$

This new threshold on $\theta$ which guarantee exponential convergence of the estimation error is significantly reduced due to the power $\frac{1}{1+j_s}$. Indeed, instead of $\mathrm{T}(\theta)$ in $L$, we have $\mathrm{T}(\theta)^{\frac{1}{1+j_s}}$.

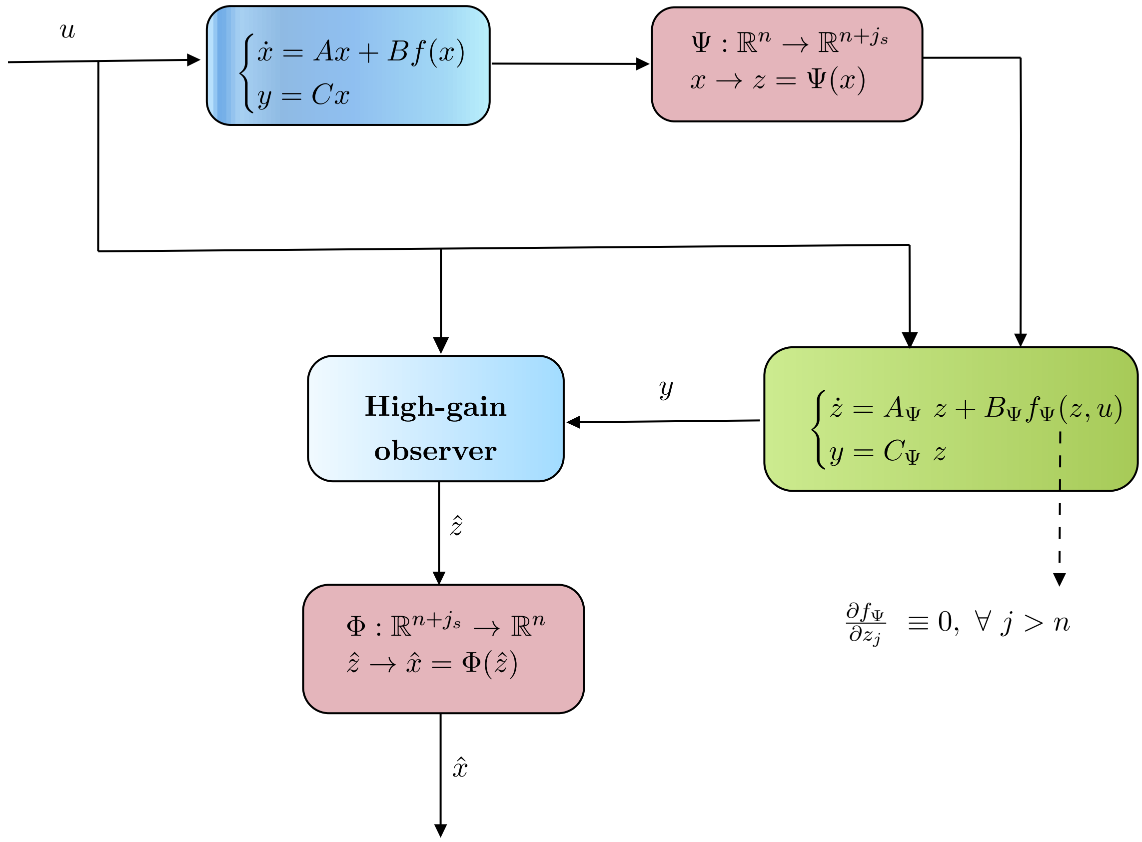



Figure 2.1: Block diagram of the high-gain observer design procedure based on system state augmentation approach.

Hence it is important to exploit condition (2.62) for systems satisfying it since it allows decreasing considerably the values of the high-gain observer. A solution is proposed in [55] by using a decomposition of the nonlinearity into two parts by using the HG/LMI technique presented in the previous section. Such a solution improves highly the standard high-gain observer, however, the decomposition of the nonlinearity affects the design of the matrices $P$ and $K$ subject to a set of $2^{j_s}$ LMIs to be solved. Hence, in this section, we propose a new design procedure technique in which we have one LMI as in the standard high-gain observer in addition to a new threshold on the design parameter $\theta$ which will result in a smaller gain as compared to the standard high-gain observer.

**System state augmentation approach**

In this section, we present the main idea of the design procedure based on the system state augmentation approach. Basically, the idea relies on transforming the original system of dimension $n$ into a new one with augmented dimension $n + j_s$, where the new nonlinear function does not depend on $j_s$ last components of the new state, we then construct a high-gain observer for the augmented system which yields a new threshold on $\theta$ attenuated to the power $\frac{1}{(1+j_s)}$. This design procedure is summarized in Figure 2.1.

The following theorem summarizes the design procedure of the proposed high-gain observer based on the system state augmentation approach.

**Theorem 2.7**

Let us consider the uniformly observable system

$$\begin{cases} \dot{x}=\psi(x,u) \\ y=\phi(x,u) \end{cases} \tag{2.65}$$

Assume there exists a state transformation (an embedding)

$$\begin{aligned} \Psi : \mathbb{R}^n &\to \mathbb{R}^{n+j_s} \\ x &\to z = \Psi(x) \end{aligned} \tag{2.66}$$

which transforms the system (2.65) into the following one

$$\begin{cases} \dot{z} = A_\Psi z + B_\Psi f_\Psi(z) \\ y = C_\Psi z \end{cases} \tag{2.67}$$

where $A_\Psi, B_\Psi$, and $C_\Psi$ have the same structure as $A, B$, and $C$, respectively, but with dimension $n + j_s$. We also have

$$f_\Psi(z) \triangleq f_\Psi(z_1, \dots, z_n) \;\Leftrightarrow\; \frac{\partial f_\Psi}{\partial z_j}(z) \equiv 0, \forall\, j > n. \tag{2.68}$$

Consider the state observer described by (2.69).

$$\begin{cases} \dot{\hat{z}} = A_\Psi \hat{x} + B_\Psi f_\Psi(\hat{z}) + L_\Psi \Big(y - C_\Psi \hat{z}\Big) \\ \hat{x} = \Phi(\hat{z}), \end{cases} \tag{2.69}$$

where $\Phi$ is a continuous left inverse of the embedding $\Psi$ satisfying $x = \Phi(z)$ and $L_\Psi \triangleq \mathrm{T}_\Psi(\theta) K_\Psi$, with $\mathrm{T}_\Psi(\theta) \triangleq \mathrm{diag}(\theta, \dots, \theta^{n+j_s})$. If there exist $P > 0$, $\lambda > 0$, $Y$, and $\theta \geq 1$ such that

$$A_\Psi^\top P + P A_\Psi - C_\Psi^\top Y - Y^\top C_\Psi + \lambda I < 0, \tag{2.70}$$

$$K_\Psi \triangleq P^{-1} Y^\top, \tag{2.71}$$

$$\theta > \theta_\Psi \triangleq \sqrt[1+j_s]{\frac{2k_{f_\Psi}\lambda_{\max}(P)}{\lambda}}, \tag{2.72}$$

then the estimation error $\tilde{x} = x - \hat{x}$ converges exponentially to zero.

*Proof.* The proof is straightforward. Indeed, from Theorem 7, if the conditions (2.70)-(2.72) are satisfied, then the error $\tilde{z} = z - \hat{z}$ converges exponentially to zero. The presence of $(1 + j_s)^{\text{th}}$ root in (2.72), as also mentioned in (2.64), is due to the fact that $f_\Psi$ does not depends on the $j_s$ last components of $z$, which leads to

$$\|\mathrm{T}_\Psi^{-1}(\theta) B_\Psi \Delta f_\Psi\| \leq \frac{k_{f_\Psi}}{\theta^{j_s}} \|\hat{\tilde{z}}\|, \tag{2.73}$$

where $\hat{\tilde{z}} = \mathrm{T}_\Psi^{-1}(\theta)\tilde{z}$. Hence, the exponential stability of $\tilde{x}$ towards zero is then preserved due to the invertibility of the mapping $\Phi$. ☐

### Particular transformation: adding a chain of integrators

This section is devoted to a special case of transforming the system (2.14) into a higher dimensional system by adding $j_s$ integrators. The aim of this section is to show that there exists a transformation satisfying the properties stated in Theorem 7.
Let us consider the following transformation

$$z = \begin{pmatrix} z_1 \\ \vdots \\ z_{n+j_s} \end{pmatrix} = \Psi(x) \triangleq \begin{pmatrix} x_1 \\ \vdots \\ x_n \\ f(x(t)) \\ \frac{\mathrm{d}f(x(t))}{\mathrm{d}t} \\ \vdots \\ \frac{\mathrm{d}^{(j_s-1)}f(x(t))}{\mathrm{d}t^{(j_s-1)}} \end{pmatrix}. \tag{2.74}$$

It is obvious to see that

$$\dot{z}_i = z_{i+1}, \text{ for } i = 1, \ldots, n+j_s-1, \tag{2.75}$$

$$\dot{z}_{n+j_s} = \frac{\mathrm{d}^{j_s}f(x(t))}{\mathrm{d}t^{j_s}} \triangleq f_\Psi(z_1, \ldots, z_n), \tag{2.76}$$

$$x = \begin{pmatrix} z_1 \\ \vdots \\ z_n \end{pmatrix} = \overbrace{\begin{bmatrix} \mathbb{I}_n & 0_{\mathbb{R}^{n\times j_s}} \end{bmatrix}}^{\Phi} z \tag{2.77}$$

where $\mathbb{I}_n$ is the identity matrix of dimension $n$.
Then following the previous section, the corresponding observer is

$$\begin{cases} \dot{\hat{z}} = A_\Psi \hat{x} + B_\Psi f_\Psi(\hat{z}) + L_\Psi\Big(y - C_\Psi \hat{z}\Big) \\ \hat{x} = \Big[\mathbb{I}_n \quad 0_{\mathbb{R}^{n\times j_s}}\Big]\, \hat{z}. \end{cases} \tag{2.78}$$

The advantage of the proposed augmentation system is the presence of the $\theta_\Psi \triangleq \sqrt[1+j_s]{\frac{2k_{f_\Psi}\lambda_{\max}(P)}{\lambda}}$, instead of $\frac{2k_{f_\Psi}\lambda_{\max}(P)}{\lambda}$ if the standard high-gain observer is applied on the augmented system. We are aware that if the standard high-gain observer is applied directly on the original system (2.14), the obtained value of $\theta_0$ in (2.24) will be smaller than $\frac{2k_{f_\Psi}\lambda_{\max}(P)}{\lambda}$. However, the presence of power $\frac{1}{1+j_s}$ will reduce significantly the values of the observer gains.

### ISS with respect to measurement noise

In this section, we compare the properties of the standard high-gain observer (2.17) and the proposed observer (2.69) with respect to measurement noises. To this end, we consider the following system where a bounded disturbance corrupts the measurement:

$$\begin{aligned} \dot{x} &= Ax + Bf(x) \\ y &= Cx + \nu \end{aligned} \tag{2.79}$$

where $\nu$ represents the disturbance affecting the measurement $y$. We will show that an upper bound on the estimation error, in an ISS sens with appropriate norms, can be ensured by the observers (2.17) and (2.69), respectively. However, we will demonstrate that using the state augmentation approach can lead to a smaller bound on the estimation error, compared to the one we get using the standard high-gain observer.

1. ISS property with Standard high-gain observer

Consider system (2.79) and the standard high-gain observer (2.17), then the transformed error dynamics system is given as

$$\dot{\hat{\tilde{x}}} = \theta \underbrace{(A - KC)}_{A_K} \hat{\tilde{x}} + \frac{1}{\theta^n} B\Delta f - K\nu. \tag{2.80}$$

Therefore, as introduced in the following proposition, the observer parameters designed by Theorem 7 ensure an ISS property.

**Proposition 2.1**

Assume that there exists a symmetric positive definite matrix $P$, a positive constant $\lambda$, and a matrix $Y$ of appropriate dimensions such that the inequalities (2.23)-(2.24) hold. Then with the observer gain $L$ given in (2.19), there exists a positive constant $\alpha$ such that the estimation error $\tilde{x}(t)$ verifies the following ISS conditions:

$$\|\tilde{x}(t)\| \leq \theta^{n-1} \sqrt{\frac{\lambda_{\max}(P)}{\lambda_{\min}(P)}} \|\tilde{x}_0\| e^{-\frac{\beta}{2}t} + \theta^n \sqrt{\frac{\gamma(1 - e^{-\beta t})}{\beta \lambda_{\min}(P)}} \sup_{s\in[0,t]} \|\nu(s)\|, \tag{2.81a}$$

$$\lim_{t\to+\infty} \|\tilde{x}(t)\| \leq \theta^n \sqrt{\frac{\gamma}{\beta \lambda_{\min}(P)}} \sup_{s\in[0,\ +\infty]} \|\nu(s)\|, \tag{2.81b}$$

where

$$\beta = \frac{\theta\lambda - 2k_f \lambda_{\max}(P) - \alpha}{\lambda_{\max}(P)}, \quad \gamma = \frac{\|Y\|^2}{\alpha}. \tag{2.82}$$

*Proof.* The stability analysis is performed using the following Lyapunov function candidate

$$V = \hat{\tilde{x}}^\top P \hat{\tilde{x}}. \tag{2.83}$$

The derivative of $V$ along the trajectory of (2.80) is given by

$$\begin{aligned} \dot{V} &= \theta\, \hat{\tilde{x}}^\top [A_K P + P A_K] \hat{\tilde{x}} + \frac{2}{\theta^n} \hat{\tilde{x}}^\top P B \Delta f - 2\hat{\tilde{x}}^\top Y^\top \nu \\ &\leq -\theta\lambda \|\hat{\tilde{x}}\|^2 + 2k_f \lambda_{\max}(P) \|\hat{\tilde{x}}\|^2 + \alpha \|\hat{\tilde{x}}\|^2 + \frac{\|Y\|^2}{\alpha} \|\nu\|^2 \\ &\leq -\underbrace{\left(\frac{\theta\lambda - 2k_f\lambda_{\max}(P) - \alpha}{\lambda_{\max}(P)}\right)}_{\beta} V + \underbrace{\frac{\|Y\|^2}{\alpha}}_{\gamma} \|\nu\|^2. \end{aligned} \tag{2.84}$$

Therefore, from the comparison theorem [117], we deduce that

$$\begin{aligned} V(t) &\leq V(0)e^{-\beta t} + \gamma e^{-\beta t}\int_0^t e^{\beta s}\|\nu(s)\|^2 ds \\ &\leq V(0)e^{-\beta t} + \gamma \sup_{s\in[0,t]} \|\nu(s)\|^2 e^{-\beta t}\int_0^t e^{\beta s} ds \\ &\leq V(0)e^{-\beta t} + \frac{\gamma}{\beta}(1-e^{-\beta t}) \sup_{s\in[0,t]} \|\nu(s)\|^2. \end{aligned} \tag{2.85}$$

Using the fact that

$$\lambda_{\min}(P)\|\hat{\tilde{x}}\|^2 \leq V(t) \leq \lambda_{\max}(P)\|\hat{\tilde{x}}\|^2,$$

we obtain

$$\begin{aligned} \|\hat{\tilde{x}}(t)\|^2 \leq \frac{\lambda_{\max}(P)}{\lambda_{\min}(P)}\|\hat{\tilde{x}}(0)\|^2 e^{-\beta t} \\ + \frac{\gamma}{\beta\lambda_{\min}(P)}(1-e^{-\beta t}) \sup_{s\in[0,t]} \|\nu(s)\|^2, \end{aligned} \tag{2.86}$$

which leads to

$$\begin{aligned} \|\hat{\tilde{x}}(t)\| \leq \sqrt{\frac{\lambda_{\max}(P)}{\lambda_{\min}(P)}}\|\hat{\tilde{x}}(0)\| e^{-\frac{\beta}{2}t} \\ + \sqrt{\frac{\gamma(1-e^{-\beta t})}{\beta\lambda_{\min}(P)}} \sup_{s\in[0,t]} \|\nu(s)\|. \end{aligned} \tag{2.87}$$

Finally, from (2.20) we get

$$\|\hat{\tilde{x}}(t)\| \leq \frac{1}{\theta}\|\tilde{x}(t)\| \quad \text{and} \quad \|\tilde{x}(t)\| \leq \theta^n \|\hat{\tilde{x}}(t)\|$$

and then the relation (2.81a) is inferred.
As for (2.81b), we use $\lim_{t\to+\infty} e^{-\beta t} = 0$ and $\lim_{t\to+\infty} \sup_{s\in[0,t]} \|\nu(s)\| = \sup_{s\in[0,\ +\infty]} \|\nu(s)\|$, which ends the proof. □

2. State augmentation approach vs standard high-gain

By a straightforward analogy, we know that we can apply the results of Proposition 1 on the augmented system defined by (2.74)-(2.77). That is the observer (2.78) designed by (2.70)-(2.72) ensures a similar ISS property than that in (2.81a)-(2.81b). However, the presence of the power $\frac{1}{1+j_s}$ in the case of augmented state-based observer (2.78), especially in the high-gain threshold condition (2.72), allows reducing significantly the values of the observer gain. For instance, for an $\epsilon > 0$, if we take $\theta = \theta_0 + \epsilon$ in the standard high-gain, then according to (2.81b), the upper bound of the estimation error satisfies:

$$\lim_{t\to+\infty} \|\tilde{x}(t)\| \leq (\theta_0 + \epsilon)^n \sqrt{\frac{\gamma}{\beta\lambda_{\min}(P)}} \sup_{s\in[0,\ +\infty]} \|\nu(s)\| \tag{2.88}$$

while with the augmented approach for $\theta = \theta_\Psi^\epsilon \triangleq \theta_\Psi + \epsilon$, we get

$$\lim_{t\to+\infty} \|\tilde{x}(t)\| \leq (\theta_\Psi^\epsilon)^{\frac{n}{1+j_s}} \sqrt{\frac{\gamma_\Psi}{\beta_\Psi\lambda_{\min}(P_\Psi)}} \sup_{s\in[0,\ +\infty]} \|\nu(s)\|. \tag{2.89}$$

It is quite clear from (2.88)-(2.89) that even for $\theta_\Psi$ greater than $\theta_0$, the presence of $\frac{n}{1+j_s}$ will reduce the effect of the measurement noise $\nu(t)$ in case of high-gain observer based on state augmentation technique. Analytically speaking, we cannot provide an explicit relation between $\theta_0$ and $\theta_\Psi$, however, from a numerical viewpoint, the value of the tuning parameter obtained by the proposed state augmentation-based observer is clearly smaller than that we get with the standard high-gain strategy. The inverse holds only if $\theta_\Psi > \theta_0^{1+j_s}$, which is hard to reach in practice.

### 2.3.5 Observer based on a combination of HG/LMI technique and system state augmentation approach

In this section, we investigate another strategy of high-gain observer design by combining the HG/LMI technique given in section 2.3.3 with the system state augmentation approach presented in section 2.3.4 to get an improved high-gain observer with more degrees on freedom as it offers the choice to select the values of two compromise indices $j_0$ and $j_s$. Furthermore, the gain of this proposed observer is significantly reduced compared to the standard high-gain observer. This gain attenuation is due to the presence of the power $\frac{n}{(1+j_s)(1+j_0)}$ in the observer tuning parameter $\theta$ which evidently results in better observer performances (less sensitivity to measurement noise and removal of the peaking).

#### Observer design

The design procedure is illustrated in Figure 2.2 which consists in transforming the original system of dimension $n$ into a system of dimension $n+j_s$, where the new nonlinear function does not depend on $j_s$ last components of the new state, we then construct a HG/LMI observer for the obtained.

Consider the system defined as

$$\begin{cases} \dot{z} = Ax + Bf(x) \\ y = Cx \end{cases} \tag{2.90}$$

Using a nonlinear transformation $\Gamma(.)$, system (2.90) is transformed into the following augmented form

$$\begin{cases} \dot{z} = A_\Gamma z + B_\Gamma f_\Gamma(z,u) \\ y = C_\Gamma z \end{cases} \tag{2.91}$$

where

$$\frac{\partial f_\Gamma}{\partial z_j}(x) \equiv 0,\ \forall\ j > n - j_s \tag{2.92}$$

As stated in the previous section, one natural solution to obtain a new system satisfying (2.92) is by adding a chain of integrators leading to the following transformation

$$z = \begin{pmatrix} z_1 \\ \vdots \\ z_{n+j_s} \end{pmatrix} = \Gamma(x) \triangleq \begin{pmatrix} x_1 \\ \vdots \\ x_n \\ f(x(t)) \\ \frac{\mathrm{d}f(x(t))}{\mathrm{d}t} \\ \vdots \\ \frac{\mathrm{d}^{(j_s-1)}f(x(t))}{\mathrm{d}t^{(j_s-1)}} \end{pmatrix}. \tag{2.93}$$

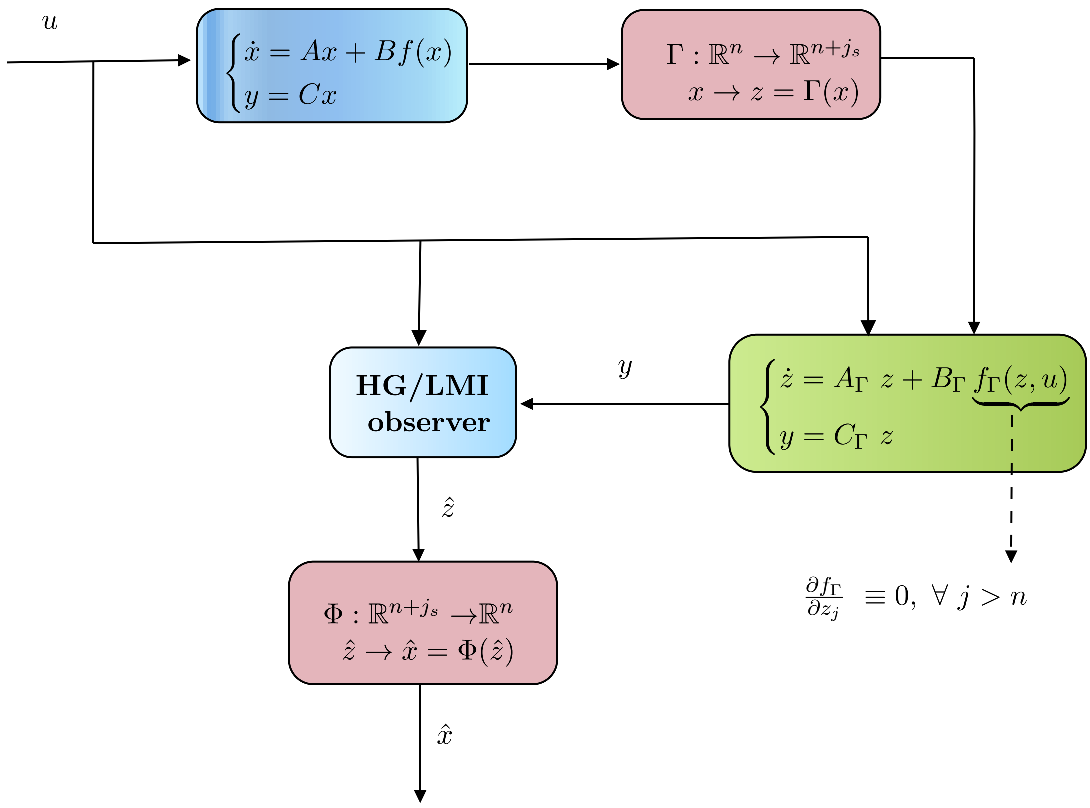



Figure 2.2: High-gain observer based on HG/LMI technique and the system state augmentation approach.

It is easy to see that $z$ obeys the following dynamics:

$$\dot{z}_i = z_{i+1}, \text{ for } i = 1, \dots, n + j_s - 1, \tag{2.94}$$

$$\dot{z}_{n+j_s} = \frac{\mathrm{d}^{j_s} f(x(t))}{\mathrm{d}t^{j_s}} \triangleq f_\Gamma(z_1, \dots, z_n), \tag{2.95}$$

$$x = \begin{pmatrix} z_1 \\ \vdots \\ z_n \end{pmatrix} = \overbrace{\begin{bmatrix} \mathbb{I}_n & 0_{\mathbb{R}^{n\times j_s}} \end{bmatrix}}^{\Phi} z \tag{2.96}$$

where $\mathbb{I}_n$ is the identity matrix of dimension $n$. The new system (2.94)-(2.95) satisfies the condition (2.92); hence, the idea consists in constructing a HG/LMI observer for system (2.94)-(2.95) and deduce an estimate $\hat{x}$ of $x$ through (2.96), namely $\hat{x} = \Phi\hat{z}$.
The corresponding observer is given as:

$$\dot{\hat{z}} = A_\Gamma \hat{z} + B_\Gamma f_\Gamma(\hat{z}) + L_\Gamma\Big(y - C_\Gamma \hat{z}\Big) \tag{2.97}$$

where $L_\Gamma \triangleq \mathrm{T}_\Gamma(\theta) K_\Gamma$, with $\mathrm{T}_\Gamma(\theta) \triangleq \mathrm{diag}(\theta, \dots, \theta^{n+j_s})$. The objective consists in determining $K_\Gamma$ and $\theta$ such that $\hat{z}$ converges exponentially to $z$. Hence, the estimated state $\hat{x}(t)$ given by

$$\hat{x}(t) = \begin{bmatrix} \mathbb{I}_n & 0_{\mathbb{R}^{n\times j_s}} \end{bmatrix} \hat{z}(t) \tag{2.98}$$

converges exponentially to $x(t)$, which is ensured by the fact that we used the HG/LMI methodology on the transformed system (2.91). For more details, we refer the reader to [55]. This is detailed in the next theorem.

**Theorem 2.8**

Assume there exist $P_\Gamma > 0$, $\lambda_\Gamma > 0$, $Y_\Gamma$, and $\sigma \geq 0$ such that the following conditions hold:

$$\mathcal{A}_\Gamma(\bar{\Psi}^\sigma)^T P_\Gamma + P_\Gamma \mathcal{A}_\Gamma(\bar{\Psi}^\sigma) - C_\Gamma^T Y_\Gamma - Y_\Gamma^T C_\Gamma + \lambda_\Gamma I < 0, \forall \bar{\Psi}^\sigma \in \mathcal{V}_{\mathcal{H}_{j_0}^\sigma}, \tag{2.99}$$

$$K_\Gamma \triangleq P_\Gamma^{-1} Y_\Gamma^\top, \tag{2.100}$$

$$\theta > \theta_\Gamma \triangleq \left(\frac{2k_{f_\Gamma}\lambda_{\max}(P_\Gamma)}{\lambda_\Gamma}\right)^{\frac{1}{(1+j_s)(1+j_0)}}, \tag{2.101}$$

where $\mathcal{H}_{j_0}^\sigma$ is as given in (2.55) with different vertices corresponding to the new nonlinear function $f_\Gamma$. Then $\hat{x}$ given by (2.98) converges exponentially to $x(t)$.

*Proof.* It is sufficient to prove that $\hat{z}$ converges exponentially to $z$, which is ensured by the fact that we use the HG/LMI methodology on the transformed system (2.91). For more details on the proof, we refer the reader to [55]. □

**Remark 2.4**

It is worth noticing that at this stage, we cannot give an explicit relation between the conditions (2.24) and (2.101). However, the presence of the power $\frac{1}{(1+j_s)(1+j_0)}$ can reduces significantly the value of the threshold in (2.101) even in case of higher values of parameters $k_{f_\Gamma}, P_\Gamma, \lambda_\Gamma$ and $j_s$.

### ISS with respect to measurement noise

By a straightforward analogy, we can obtain the ISS property of the observer proposed in this section which is based on the combination of the HG/LMI technique and the state augmentation approach.
To this end, we consider the following system where the measurement is corrupted by a bounded disturbance:

$$\begin{array}{rcl} \dot{x} &=& Ax + Bf(x) \\ y &=& Cx + \nu \end{array} \tag{2.102}$$

where $\nu$ represents the disturbance affecting the measurement $y$.
The transformed error dynamics system is given as

$$\dot{\tilde{\bar{x}}} = \theta \underbrace{(A - KC)}_{A_K} \tilde{\bar{x}} + \frac{1}{\theta^n} B\Delta f - K\nu. \tag{2.103}$$

Therefore, the observer parameters designed by Theorem 8 ensure an ISS property as introduced in the following proposition.

**Proposition 2.2**

Assume that there exists a symmetric positive definite matrix $P_\Gamma$, a positive constant $\lambda$, and a matrix $Y$ of appropriate dimensions such that the inequalities (2.99)-(2.101) hold. Then with the observer gain $L$ given in (2.19), there exists a positive constant $\alpha$ such that the estimation error $\tilde{x}(t)$ verifies the following ISS conditions:

$$\|\tilde{x}(t)\| \leq \theta_\Gamma^{\frac{n-1}{(1+j_s)(1+j_0)}} \sqrt{\frac{\lambda_{\max}(P_\Gamma)}{\lambda_{\min}(P_\Gamma)}} \|\tilde{x}_0\| e^{-\frac{\beta_\Gamma}{2}t} + \theta_\Gamma^{\frac{n}{(1+j_s)(j+j_0)}} \sqrt{\frac{\gamma(1-e^{-\beta_\Gamma t})}{\beta \lambda_{\min}(P_\Gamma)}} \sup_{s\in[0,t]} \|\nu(s)\|, \tag{2.104a}$$

$$\lim_{t\to+\infty} \|\tilde{x}(t)\| \leq (\theta_\Gamma^\epsilon)^{\frac{n+j_s}{(1+j_s)(1+j_0)}} \sqrt{\frac{\gamma_\Gamma}{\beta_\Gamma \lambda_{\min}(P_\Gamma)}} \sup_{s\in[0,\ +\infty]} \|\nu(s)\|. \tag{2.104b}$$

where

$$\beta_\Gamma = \frac{\theta_\Gamma \lambda - 2k_f \lambda_{\max}(P_\Gamma) - \alpha}{\lambda_{\max}(P_\Gamma)},\ \gamma = \frac{\|Y\|^2}{\alpha}. \tag{2.105}$$

**Remark 2.5**

The advantage of the observer proposed in this section lies in its ability to combine the classical standard high-gain observer, the HG/LMI-based observer, and the observer based on the system state augmentation approach. This allows the observer to benefit simultaneously from the advantages of each of these three methodologies, specifically, the observer is able to avoid the peaking phenomenon, reduce the sensitivity to high-frequency measurement noise, and enhance the convergence rate if necessary. To get a good trade-off between all these criteria, the new observer offers the possibility to play with the values of $j_0$ and $j_s$ and in the case of $j_s = j_0 = 0$, we have the particular case of the standard high-gain observer.

## 2.4 Conclusion

This chapter has been devoted to the design of nonlinear observers which relies on high-gain techniques. A design strategy based on the system state augmentation approach is proposed then a combination of this latter with the HG/LMI technique from the work in [55] is established. The main motivation of the new structure is overcoming the drawbacks which make the use of high-gain observers questionable in practical applications. It was shown that by augmenting the state of the system under consideration we may reduce the value of the tuning parameter $\theta$ and thereafter, the value of the observer gain is attenuated.
Another design procedure proposed in this chapter is the combination of the system state augmentation approach together with the HG/LMI technique. This combination allows to benefit from the advantages of each methodology, hence, thanks to the HG/LMI technique which allows a particular decomposition of the nonlinearity, the value of the Lipschitz constant is reduced resulting in a smaller value of the tuning parameter and consequently an even smaller gain.
Since the main problem of the high-gain observer is the amplification of high-frequency measurement noise, we have considered the case where the measurement is affected by noises, then we have investigated the stability and robustness properties of the considered high-gain techniques

using the theory of input-to-state stability. Moreover, we have shown that the ISS properties of the estimation error dynamics from measurement noise are not deteriorated as compared with the classical high-gain observer case.

Chapter

3

# Moving horizon estimator for quasi-LPV systems

*"Study the past, if you would divine the future."*

Confucius

## Contents



## 3.1 Introduction

It is well established that the Kalman filter is the optimal state estimator for unconstrained, linear systems subject to normally distributed state and measurement noise. Many physical systems, however, exhibit nonlinear dynamics and have states subject to hard constraints, hence, Kalman filtering is no longer directly applicable. As a result, many different types of nonlinear state estimators have been proposed, such as extended Kalman filters, moving horizon estimation, model inversion, and Bayesian estimation. The extended Kalman filter appears to be the standard technique that is used in a number of nonlinear estimation applications due to its relative simplicity and demonstrated efficacy in handling nonlinear systems. However, the extended Kalman filter, or EKF, is at best an ad hoc solution to a difficult problem, and hence there exist many barriers to the practical implementation of EKFs. Some of these problems include the inability to accurately incorporate physical state constraints and poor use of the nonlinear model. This was one of the main motivations for early ventures into moving horizon estimation techniques.

MHE is an online optimization strategy that accurately employs the nonlinear model and incorporates information about the system in the form of constraints to improve the estimate. First ideas on MHE date back to the sixties [125], motivated by its intrinsic robustness. In the moving horizon estimation approach, the state estimate is determined by solving an optimization problem online that minimizes the sum of the squared errors. At a sampling time, when a new measurement is obtainable, the older measuring within the estimation window is discarded and the horizon finite optimization problem is solved again to obtain the new state estimate [126–128]. The cost function is usually made up of two contributions: a prediction error computed on a recent batch of inputs and outputs; an arrival cost that serves the purpose of summarizing the past data.

Moving horizon estimation techniques gained increasing interest over the last decades and were explicitly addressed first in the papers [58,129,130]. Thereafter, researchers have focused on the application of such techniques to linear systems [115,126,131,132], hybrid systems [133,134], and nonlinear systems [58, 59, 130, 135–137]. In [59] an asymptotic state observer is described that results from the numerical solution of a sequence of nonlinear algebraic equations via Newton's method. Similar optimization-based solution techniques are employed in [130, 135] to construct stable estimators for continuous-time dynamic systems. In [58], a moving horizon estimator for nonlinear continuous-time systems was proposed that performs estimation at discrete time instants by approximately minimizing an integral error defined on the preceding time window. In [136], a moving horizon estimation scheme was presented that allows one to explicitly take into account possible constraints on the system and requires the solution of a nonlinear programming problem at each time step. Moreover, a sufficient condition for the non-divergence of the estimation error in the presence of bounded noises was provided. In [138], system uncertainties have been explicitly considered and in [139]- [140] distribution and decentralization have been investigated.

In this third part of the manuscript, we address the problem of state estimation of nonlinear plants rewritten in the form of quasi-LPV discrete-time systems using the moving horizon estimation approach. In general, the problem of estimation of LPV systems consists in fitting the model to a set of measurements. A parameter estimator calculates the parameter values that render the model-predicted values of process outputs closest to the corresponding measured values of the process outputs. In offline parameter estimation, a model is fitted optimally to the process measurements from one or several completed process runs [141]. However, in online parameter estimation, a model is fitted optimally to the past and present process measurements while the process is in operation [142–144]. The available methods on parameter estimation via state estimation include extended Kalman filter [145–149], reduced-order observer [150, 151], proportional observer [152, 153], interval observers [154–157], generalized dynamic observer [158, 159], adaptive observer [160–162].

The main contribution of this chapter is the design and analysis of a robust estimator for nonlinear systems under bounded disturbances combining quasi-LPV models and dual estimation using a receding horizon framework. The proposed algorithm simultaneously estimates the unknown parameters and the states using a dual estimation approach by focusing on two different formulations, which will be referred to as "optimistic" and "pessimistic" MHE. The conditions to guarantee robust stability and convergence to the true system and states in the presence of bounded disturbances are derived. To achieve these results the prior weighting in the cost function and the length of the estimation horizon are properly chosen and the optimization is performed within a multiple iteration scheme that improves the performance of the estimation at each sample.

## 3.2 Optimization based estimation techniques

Basically, there are two classes of state estimators in systems and control theory. The first class of state estimators is the class of estimators to which belong for example the Luenberger observer. In this traditional class, state estimators are realized in a typically system-theoretic fashion, namely as control systems where the available measurements are fed into the input of the estimator and the output delivers the estimated state. In this class of estimation techniques, an explicit observer equation must be derived such that the induced dynamics on the estimation error is provably asymptotically stable. However, the search for such an observer candidate is clearly a hard task as a high level of genericity is required. Contrary to analytic observers that use the explicit study of the state estimation error in order to design the observer correction term, optimization-based observers use the very definition of observability in order to derive the state estimation algorithm. The idea is to use the fact that as long as the nominal system is considered, estimating the state of an observable system is equivalent to minimizing a cost function which is usually taken as the sum of the squared output prediction error over some observation horizon. Consequently, if an algorithm can guarantee that this quantity converges asymptotically to 0, then there is no need for additional proof of convergence. The convergence of the state estimation error is a direct consequence of the convergence of the cost function [163]. In this section, an overview of optimization-based estimation techniques is presented briefly.

### 3.2.1 Optimal linear state estimation

The following discrete-time linear system is considered

$$\begin{aligned} x_{k+1} &= Ax_k + Bu_k + w_k, \\ y_k &= Cx_k + v_k, \end{aligned} \tag{3.1}$$

where $x \in \mathbb{R}^n$ is the state vector, $u \in \mathbb{R}^m$ is the input, $y \in \mathbb{R}^p$ is the output or measurement vector, $w \in \mathbb{R}^n$ is the process disturbance vector and $v \in \mathbb{R}^p$ is the measurement noise vector.
This model can be derived from a continuous-time system in which the output measurements are available at equally spaced sampling times, the input remains constant over this sampling period, and the output is sampled at the same time the input is injected. We assume that the input sequence is bounded and known exactly and that a *priori* estimate of the initial state, $\bar{x}_0$, is available.
For the stochastic linear system in (3.1), it can be shown that when $w_k$ and $v_k$ are independent, zero mean, normally distributed random variables with covariances $Q$ and $R$, respectively, and $\bar{x}_0$ is an independent, normally distributed random variable with covariance $Q_0$, the Kalman

filter produces the optimal estimate of the state [89]. For linear Gaussian systems, this estimate is also the maximum likelihood estimate. The Kalman filter recursively estimates the state of the linear system given by (3.1) from the output measurements at each sampling time as follows

- Time update/ prediction step/ a-priori:

  prediction of the state

$$\hat{x}_{k|k} = A\hat{x}_{k-1|k-1} + Bu_{k-1}, \tag{3.2}$$

  projection of the error covariance

$$P_{k|k-1} = AP_{k-1}A^T + Q, \tag{3.3}$$

- Measurement update/ correction step/ a-posteriori

  computation of the Kalman gain

$$K_k = P_{k|k-1}C^T\left(CP_{k|k-1}C^T + R\right)^{-1}, \tag{3.4}$$

  update of the estimate with the measurement

$$\hat{x}_{k|k} = \hat{x}_{k|k-1} + K_k\left(y_k - C\hat{x}_{k|k-1}\right), \tag{3.5}$$

  update of the error covariance

$$P_k = (I - K_kC)\,P_{k|k-1}, \tag{3.6}$$

where $\hat{x}_{k|k}$ is the optimal estimate of the state at sample time $k$ given $k$ output measurements, $Q$ and $R$ present the measure of confidence in the model and the measurement, respectively. If the covariance of the measurement noise, $R$, is small relative to the covariance of the process noise, $Q$, then the measurements are relatively noise-free and the deviations between the measured output and the predicted output should be made small. If the measurement noise covariance is large relative to the process noise covariance, then the measurements are relatively noisy and the feedback correction to the model prediction should be small.
As given by (3.2) and (3.5), the Kalman filter follows a two-step procedure to calculate the maximum a-posteriori Bayesian estimate [164]. The first step is the time update, the second step is the measurement update. The first step uses the system model to predict the current state of the system based on the last estimate. In the measurement step, this prediction is updated by using the measurements. The Kalman filter is hence a predictor-corrector type estimator that is optimal in the sense that it minimizes the estimated error covariance. The filter gain is computed at each sample time from the expression in (3.4) in which $P_k$ is the covariance of the state estimate. The solution of the discrete algebraic Riccati equation

$$P = A^TPA - A^TPB\left(B^TPB + R\right)^{-1}B^TPA + Q \tag{3.7}$$

can be used for the calculation of the steady-state gain which makes $_{Pk+1} = Pk = P$ constant. The stability of the Kalman filter is ensured provided that $R$ is positive definite, $Q$ and $Q_0$ are positive semidefinite, and the system (3.1) is detectable. Then the reconstruction error, $ek = x_k - \hat{x}_{k|k-1}$, converges to zero for the nominal system with no state or measurement noise.

### 3.2.2 Extended Kalman filter

A straightforward approximation to optimal nonlinear state estimation is to linearize the nonlinear model about a given operating point and apply optimal linear state estimation to the linearized system which gives rise to the extended Kalman filter (EKF). hence, the EKF is just an extension of the Kalman filter to nonlinear systems. This means that the difference with the EKF is that the state and/or the output equations can contain nonlinear functions. So rather than considering systems of the form (3.1), the EKF can consider a more general nonlinear set of equations:

$$x_k = f(x_{k-1}, u_{k-1}) + w_{k-1} \tag{3.8}$$

$$y_k = h(x_k) + v_k \tag{3.9}$$

where $f$ is the vector-valued nonlinear state transition function and $h$ is the vector-valued, nonlinear observation or output function. A linearized model of the system about the state $x^*$ can be developed from the Taylor series expansion. Depending on the selection of $x^*$, several variations of the extended Kalman filter can be developed. The most common approach is the first-order filter in which the nonlinear system is linearized about the current state estimate at each sampling time using the first-order terms of the Taylor expansion. Further discussion of extended Kalman filtering is contained in [89, 165–167].
We summarize the algorithm for implementing the EKF presented in [167],

$$\hat{\mathbf{x}}_{k|k-1} = \mathbf{f}\left(\hat{\mathbf{x}}_{k-1}, \mathbf{u}_{k-1}\right) \tag{3.10}$$

$$\mathbf{P}_{k|k-1} = \mathbf{F}_{k-1}\mathbf{P}_{k-1}\mathbf{F}_{k-1}^T + \mathbf{Q}_{k-1} \tag{3.11}$$

$$\mathbf{K}_k = \mathbf{P}_{k|k-1}\mathbf{H}_k^T\left(\mathbf{H}_k\mathbf{P}_{k|k-1}\mathbf{H}_k^T + \mathbf{R}_k\right)^{-1} \tag{3.12}$$

$$\hat{\mathbf{x}}_k = \hat{\mathbf{x}}_{k|k-1} + \mathbf{K}_k\left[\mathbf{z}_k - \mathbf{h}\left(\hat{\mathbf{x}}_{k|k-1}\right)\right] \tag{3.13}$$

$$\mathbf{P}_k = \left(\mathbf{I} - \mathbf{K}_k\mathbf{H}_k\right)\mathbf{P}_{k|k-1} \tag{3.14}$$

where

$$F_k = \left.\frac{\partial f(x,t)}{\partial x}\right|_{\hat{x}_{k-1}} \tag{3.15}$$

$$H_k = \left.\frac{\partial h(x,t)}{\partial x}\right|_{\hat{x}_{k|k-1}} \tag{3.16}$$

are the corresponding Jacobian matrices of $f$ and $h$, respectively.
An important feature of the EKF is that the Jacobian $H_k$ in the equation for the Kalman gain $K_k$ serves to correctly propagate or "magnify" only the relevant component of the measurement information. For example, if there is not a one-to-one mapping between the measurement and the state via $h$, the Jacobian $H_k$ affects the Kalman gain so that it only magnifies the portion of the residual $z_k - h(\hat{x}_k^-, 0)$ that does affect the state. If overall measurements, there is not a one-to-one mapping between the measurement $z_k$ and the state via $h$, then as you might expect the filter will quickly diverge. In this case, the process is unobservable. A block diagram of the extended Kalman filter is given in Figure 3.1.

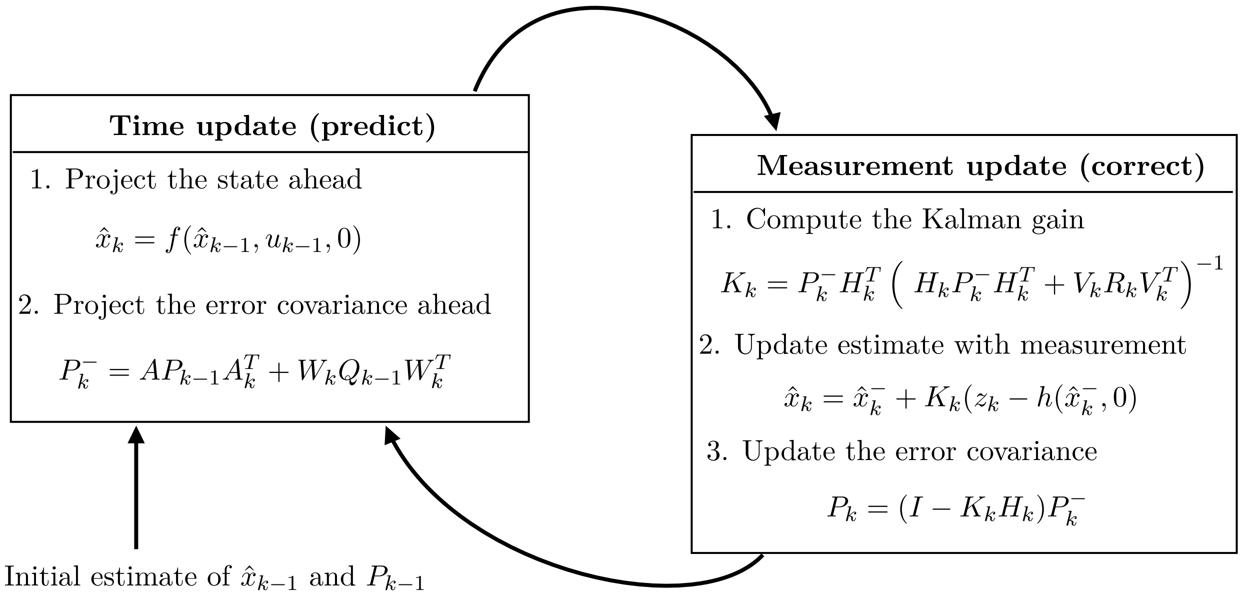



Figure 3.1: A complete picture of the operation of the extended Kalman filter.

### 3.2.3 Linear batch state estimation

If the assumption on $w_k$ and $v_k$ does not hold anymore, i.e., they are not Gaussian random variables which can be considered as process and measurement disturbances with unknown statistics, then, an optimal state estimate in the probabilistic sense cannot be obtained. However, an estimator that provides the best state estimate can be established in a deterministic sense based on the following least squares problem.

$$\begin{aligned}\min_{\{\hat{w}_{-1|k},\ldots,\hat{w}_{k-1|k}\}} \quad & \Phi_k = \hat{w}_{-1|k}^T Q_0^{-1} \hat{w}_{-1|k} \\ & + \sum_{j=0}^{k-1} \hat{w}_{j|k}^T Q^{-1} \hat{w}_{j|k} + \sum_{j=0}^{k} \hat{v}_{j|k}^T R^{-1} \hat{v}_{j|k}\end{aligned} \tag{3.17}$$

subject to:

$$\hat{x}_{0|k} = \bar{x}_0 + \hat{w}_{-1|k} \tag{3.18}$$

$$\hat{x}_{j+1|k} = A\hat{x}_{j|k} + Bu_j + \hat{w}_{j|k} \tag{3.19}$$

$$y_j = C\hat{x}_{j|k} + \hat{v}_{j|k} \tag{3.20}$$

where $\hat{x}_{0|k}$ is the estimate of $x_0$ given $k$ output measurements, $\bar{x}_0$ is an a priori estimate of the initial state, $\hat{w}_{j|k}$ are the estimated process disturbances, and $\hat{v}_{j|k}$ are the estimated measurement disturbances. The weighting matrices, $Q_0^{-1}$, $Q^{-1}$, and $R^{-1}$, specify the relative contribution of each of the terms in the quadratic objective and are the tuning parameters for the least squares estimator. This approach attempts to minimize the estimated process and measurement disturbances in the least squares sense. The choice of the weighting matrices is based on a compromise between minimizing the estimated process disturbances and minimizing the estimated measurement disturbances in a manner similar to the measurement and state noise covariance matrices used in the Kalman filter.

The solution to the least squares problem, $\hat{w}^*_{j|k}$ is used to compute the state estimate at time $j$ given $k$ output measurements, $\hat{x}_{j|k}$, as follows

$$\hat{x}_{j|k} = A^j \bar{x}_0 + \sum_{i=0}^{j} A^{j-i} \hat{w}^*_{i-1|k} + \sum_{i=1}^{j} A^{j-i} B u_{i-1} \tag{3.21}$$

**Remark 3.1**

This expression computes a smoothed state estimate when $j < k$, a filtered state estimate when $j = k$, and a predicted state estimate when $j > k$. It can be shown that the filtered estimate from the batch least squares estimator is the optimal filtered estimate for the stochastic system in (3.1) when $v_k$, $w_k$ and $x_0$ follow the same assumptions made in the Kalman filter and the weighting matrices are chosen as the inverse of the corresponding covariance matrices [89].

The choice of the coefficients of the weighting matrices is a compromise between minimizing the estimated process disturbances versus minimizing the estimated measurement disturbances. This choice is based on the expected magnitudes of each of these disturbances. If the output measurements are reliable, then $R^{-1}$ is chosen to be large relative to $Q^{-1}$. On the other hand, if the output measurements are poor, then $R^{-1}$ is chosen to be small relative to $Q^{-1}$. In this case, the model is assumed to be a more reliable indication of the state than the output measurements.

### 3.2.4 Nonlinear batch state estimation

We now consider the following nonlinear discrete system model

$$\begin{aligned} x_{k+1} &= f(x_k, u_k) + w_k, \\ y_k &= h(x_k) + v_k, \end{aligned} \tag{3.22}$$

where $f : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^n$ and $h : \mathbb{R}^n \to \mathbb{R}^p$ are continuous functions and as in the linear case, the input sequence $u_k$ is assumed to be bounded and known and that a *priori* estimate of the initial state, $x_0$, is available.

Consider any admissible input sequence $u_0, u_1, \ldots, u_N$ and two state values, $x_0, z_0$. Let $x_j$ and $z_j$, $j = 1, \ldots, N$ denote the state sequences generated by the model, $x_j = f(x_{j-1}, u_{j-1})$ and $z_j = f(z_{j-1}, u_{j-1})$. We require the following observability condition for the nonlinear system. For any two states, $x_o$ and $z_o$, and all admissible input sequences, there exists a horizon length $N \in [n-1, \infty)$ and a function $\alpha : \mathbb{R}^+ \to \mathbb{R}^+$ such that

$$\sum_{j=0}^{N} \| g(z_j) - g(x_j) \| \geq \alpha \left( \| z_0 - x_0 \| \right) \tag{3.23}$$

where $\alpha$ is a continuous, increasing function with $\alpha(0) = 0$. This condition is similar to the uniform observability property in [168] referred to as Property $O$. It is a weaker discrete-time

version of that presented in [58] for continuous systems and a stronger condition than strong observability or finite time observability presented in [74].
The nonlinear batch state estimator is formed in a similar manner to the linear batch state estimator presented previously

$$\min_{\{\hat{w}_{-1|k},\ldots,\hat{w}_{k-1|k}\}} \Phi_k = \hat{w}_{-1|k}^T Q_0^{-1} \hat{w}_{-1|k} + \sum_{j=0}^{k-1} \hat{w}_{j|k}^T Q^{-1} \hat{w}_{j|k} + \sum_{j=0}^{k} \hat{v}_{j|k}^T R^{-1} \hat{v}_{j|k} \tag{3.24}$$

subject to:

$$\hat{x}_{0|k} = \bar{x}_0 + \hat{w}_{-1|k} \tag{3.25}$$

$$\hat{x}_{j+1|k} = f(\hat{x}_{j|k}, u_k) + \hat{w}_{j|k} \tag{3.26}$$

$$y_j = g(\hat{x}_{j|k}) + \hat{v}_{j|k} \tag{3.27}$$

The solution to the least squares problem, $\hat{w}_{j|k}^*$ is used to compute the state estimate at time $j$ given $k$ output measurements, $\hat{x}_{j|k}$, as follows

$$\hat{x}_{j+1|k} = f(\hat{x}_{j|k}, u_k) + \hat{w}_{j|k}^* \tag{3.28}$$

$$\hat{x}_{0|k} = \bar{x}_0 + \hat{w}_{-1|k}^* \tag{3.29}$$

Convergence of the constrained batch state estimator for linear systems is proved in [169]. A convergence proof for the nonlinear system in (3.22) is given in [109]. It was shown that the filtered state estimate converges to the true state for a non-zero initial reconstruction error and no model mismatch. The relation (3.30) is obtained from optimization at each time $k$

$$\phi_k - \hat{v}_{k|k}^{*T} R^{-1} \hat{v}_{k|k}^* - \hat{w}_{k-1|k}^{*T} Q^{-1} \hat{w}_{k-1|k}^* \geq \phi_{k-1} \tag{3.30}$$

Since $Q$ and $R$ are positive definite, the sequence $\phi_k$ is monotonically non-decreasing and bounded above and therefore converges. The convergence of $\phi_k$ and equation (3.30) imply $\hat{w}_{k-1|k}^*$ and $\hat{v}_{k-1|k}^*$ converge to zero. From the continuity of $f$ and the observability condition given in (3.23), it follows that $\hat{x}_{k|k}$ converges to $x_k$.

## 3.3 Moving horizon state estimation

One of the main problems of the batch state estimation seen in the previous section is that it becomes intractable as time goes on. Hence, the moving horizon estimator (MHE) arose to face this drawback. In this approach, the state is estimated from the most recent $N + 1$ output measurements. These measurements are referred to as the history or observer horizon of length $N$. The horizon can be viewed as a window of past output measurements that moves forward in time at each sampling time when a new measurement is available. In this section, a summary of the moving horizon estimation scheme is provided. Most of the theoretical background can be found in detail in [112] and references therein.

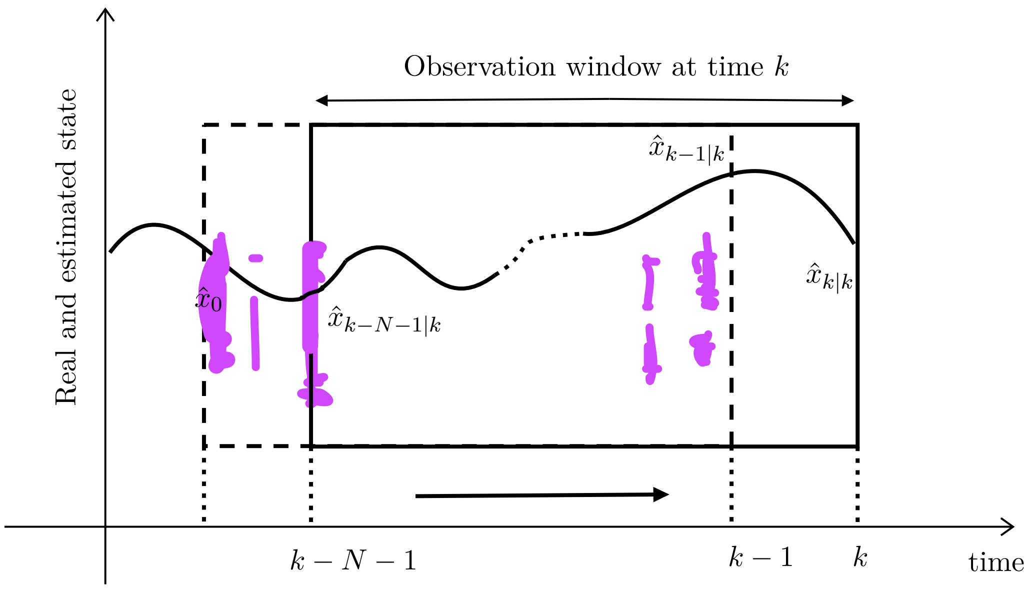



Figure 3.2: Concept of moving horizon estimation approach.

### 3.3.1 Moving horizon for linear systems

A recursive form of the linear batch state estimation problem in (3.17) can be constructed for system (3.1) using the moving horizon approach. In this approach, the state estimate at time $k$ is determined recursively from the solution of a least squares problem that uses the predicted estimate at time $k-N-1$ denoted by $\hat{x}_{k-N|k-N-1}$ and the most recent $N+1$ measurements. The least squares problem is defined as follows

$$\begin{aligned}\min_{\{\hat{w}_{k-N-1|k},\ldots,\hat{w}_{k-1|k}\}} \Phi_k &= \hat{w}_{k-N-1|k}^T P_{k-N}^{-1} \hat{w}_{k-N-1|k} \\ &+ \sum_{j=k-N}^{k-1} \hat{w}_{j|k}^T Q^{-1} \hat{w}_{j|k} + \sum_{j=k-N}^{k} \hat{v}_{j|k}^T R^{-1} \hat{v}_{j|k}\end{aligned} \tag{3.31}$$

subject to:

$$\hat{x}_{k-N|k} = \hat{x}_{k-N|k-N-1} + \hat{w}_{k-N-1|k} \tag{3.32}$$

$$\hat{x}_{j+1|k} = A\hat{x}_{j|k} + Bu_j + \hat{w}_{j|k} \tag{3.33}$$

$$y_j = C\hat{x}_{j|k} + \hat{v}_{j|k} \tag{3.34}$$

The moving horizon allows for a finite number of decision variables at each sampling time. The filtered state estimate at time $k$ is determined recursively from the predicted estimate at time $k-N-1$ and the most recent $N+1$ output measurements. The first $N$ estimates are computed using the batch estimator to initialize the observer horizon. The state estimate at time $k-N+j$ given $k$ output measurements denoted by $x_{k-N+j|k}$, is computed from the solution of the least squares problem in a manner similar to the batch estimator.

$$\hat{x}_{k-N+j|k} = A^j \hat{x}_{k-N|k-N-1} + \sum_{i=0}^{j} A^{j-i} \hat{w}_{k-N-1+i|k}^* + \sum_{i=1}^{j} A^{j-i} B u_{k-N-1+i} \tag{3.35}$$

**Remark 3.2**

This expression computes a smoothed state estimate when $j < N$, a filtered state estimate when $j = N$, and a predicted state estimate when $j > N$. It can be shown that the filtered estimate from the moving horizon estimator is the optimal filtered estimate for the stochastic system in (3.1) in which $v_k$, $w_k$ and $x_0$ follow the same assumptions made in the Kalman filter and the weighting matrices $P_{k-N}^{-1}$, $Q^{-1}$ and $R^{-1}$, are the inverses of the nonsingular covariance matrices for $\hat{x}_{k-N|k-N-1}$, $w_k$ and $v_k$, respectively [169]. The matrix $P_{k-N}$ is the discrete filtering Riccati matrix at time $k - N$ and is computed using the recursion in (3.7).

### MHE constraints

One of the main advantages of the MHE-based schemes over the Kalman filter is the direct constraints handling. The addition of constraints to the estimator results in a deterministic quadratic programming problem that attempts to minimize the output prediction errors and state disturbances subject to the constraints which can be expressed as follows

$$h_{min} \leq H\hat{x}j \mid k \leq h_{max} \tag{3.36}$$

$$w_{min} \leq \hat{w}i \mid k \leq w_{max} \tag{3.37}$$

The estimated state constraints in (3.36) specify maximum and minimum limits on the state estimates. These constraints are applied to prevent physically unrealistic state estimates and the estimated state disturbance constraints in (3.37) specify an upper and lower bound on the estimated state disturbances. These constraints can be viewed as altering the distribution of the state disturbances such that the probability of a state disturbance outside of the constraints is zero. These constraints prevent estimated state disturbances that cannot realistically occur in the process.

**Remark 3.3**

The constraints in (3.36)-(3.37) are imposed based on a heuristic argument with no probabilistic justification. Therefore, this estimator is not optimal in any probabilistic sense even for linear, Gaussian systems. However, the constraints allow for the implementation of a reasonably simple estimator that can handle a complex, constrained stochastic system without detailed probabilistic analysis [109].

## 3.3.2 Moving horizon for nonlinear systems

Generalizing ideas from linear filtering [125], early formulations of nonlinear moving horizon estimator were developed in [58,129,130,170–172]. A direct approach to the deterministic discrete-time nonlinear MHE problem is to view the problem as one of inverting a sequence of nonlinear algebraic equations defined from the state update and measurement equations, and some moving time horizon [59] Moraal. Such discrete-time observers are formulated in the context of numerical nonlinear optimization and analyzed with respect to convergence in [**?**,136,173–175]. In recent contributions [176] provides results on how to use a continuous time model in the

discrete-time design, while issues related to parameterization are highlighted in [163] computational efficiency are central targets of [163, 174, 177]. Uniform observability is a key assumption in most formulations and analyses of nonlinear MHE. For many practical problems, like combined state and parameter estimation problems, uniform observability is often not fulfilled and modifications are needed to achieve robustness [59, 178].
In this section, the moving horizon state estimator presented in the previous section for linear systems is extended to nonlinear systems. The main objective is to define a deterministic least squares estimator for the discrete nonlinear system given in (3.22).

### Estimated state disturbance approach

Since it is not possible to obtain a general closed-form recursive solution to the nonlinear batch state estimation problem as stated before, it is also difficult to develop a moving horizon estimator equivalent to the nonlinear batch estimation problem. Therefore, a moving horizon estimator similar to the linear case is usually constructed as follows

$$\min_{\{\hat{w}_{k-N-1|k},\dots,\hat{w}_{k-1|k}\}} \Phi_k^N = + \sum_{j=k-N}^{k-1} \hat{w}_{j|k}^T Q^{-1} \hat{w}_{j|k} + \sum_{j=k-N}^{k} \hat{v}_{j|k}^T R^{-1} \hat{v}_{j|k} \tag{3.38}$$

subject to:

$$\hat{x}_{j+1|k} = f\left(\hat{x}_{j|k}, u_j\right) + \hat{w}_{j|k} \tag{3.39}$$

$$\hat{v}_{j|k} = y_j - g\left(\hat{x}_{j|k}\right) \tag{3.40}$$

This nonlinear moving horizon estimator does not penalize the initial state disturbances in the horizon allowing the initial state estimate $\hat{x}_{k-N|k}$ in the horizon to be chosen freely. The advantage of this approach is that nominal convergence of the estimated state to the true state is achieved [109].
Since the process model is nonlinear, the MHE technique requires the solution to a general nonlinear optimization problem. The state estimate at sample time $n+j$ where $n = k - N$, is computed from the optimal solution using the following recursion.

$$\begin{aligned} \hat{x}_{n+j+1|k} &= f\left(\hat{x}_{n+j|k}, u_{n+j}\right) + \hat{w}_{n+j}^* \\ \hat{x}_{n|k} &= \hat{x}_{n|k}^* \end{aligned} \tag{3.41}$$

### Initial State Estimate Approach

Another approach is based on estimating the initial state without including the estimated state disturbances in the least squares problem. This approach is well motivated if there are no state or process disturbances and that the measurement is corrupted by zero-mean noise [165]. Hence, the objective is to determine the initial estimate that minimizes the difference between the measured output and the predicted output throughout the horizon.

$$\min_{\{\hat{w}_{k-N-1|k},\dots,\hat{w}_{k-1|k}\}} \Phi_k^N = + \sum_{j=k-N}^{k-1} \hat{v}_{j|k}^T R^{-1} \hat{v}_{j|k} \tag{3.42}$$

subject to:

$$\hat{x}_{j+1|k} = f\left(\hat{x}_{j|k}, u_j\right) \tag{3.43}$$

$$y_j = g\left(\hat{x}_{j|k}\right) + \hat{v}_{j|k} \tag{3.44}$$

The nominal convergence of this approach for continuous-time systems is shown in [58]. The proof of convergence for discrete systems follows in the same manner as the moving horizon presented in the previous section. Continuous-time systems with discrete time measurements can also be considered with this approach.

The advantage of this initial state estimate approach is a smaller number of decision variables for a given horizon length which reduces the complexity of the optimization problem since the horizon length is the only tuning parameter. However, neglecting the disturbances is not realistic for many applications. issues such as unmeasured or unmodeled disturbances, process modeling errors, and variation in the model parameters cannot be addressed adequately when only the initial state horizon is estimated.

## 3.4 Moving horizon state estimation for quasi-LPV systems

The assumptions of constant model parameters may not always be realistic in many practical applications in which parameters may change during the operation. In this situation, a more accurate determination of the state can be made by allowing one or more model parameters to vary. The parameter values are then estimated along with the state which is referred to as combined state and parameter estimation. In this section, we are going to address the problem of MHE design on LPV systems written under the form of quasi-LPV formulation.

Combined state and parameter estimation typically is performed by augmenting the state of the system to include the parameters to be estimated simultaneously with the state. The most common assumption made to describe the dynamic behavior of these parameters is an integrated white noise process. Let's consider the discrete-time system

$$\begin{aligned} x_{k+1} &= f(x_k, u_k) \\ y_k &= g(x_k) \end{aligned} \tag{3.45}$$

Using the assumption given before, (3.45) is augmented as follows.

$$\begin{aligned} \begin{bmatrix} x_{k+1} \\ \Theta_{k+1} \end{bmatrix} &= \begin{bmatrix} f\left(x_k, u_k, \Theta_k, k\right) \\ \Theta_k \end{bmatrix} \\ y_k &= g\left(x_k, \Theta_k, k\right) \end{aligned} \tag{3.46}$$

where $\Theta$ is the vector of parameters to be estimated.

The observability of the augmented system is not the same as the observability of the original system, hence it should be verified using the observability techniques seen in the first chapter.

Parameter estimation using the extended Kalman filter is discussed in [179, 180]. The extended Schmidtâ€" Kalman filter is discussed in Jazwinski [89]. This filter accounts for the effect of parameter variation on the state but does not compute an estimate of the parameter values. A discussion of modeling techniques that can improve the estimation of parameter values is presented in [181]. The authors in [182] recommend a decoupled estimator in which the states and parameters are computed separately to reduce the nonlinearity due to parameter-state interactions.

### 3.4.1 LPV modeling

Linear parameter varying (LPV) systems are described by linear differential equations whose parameters depend on some online measured signals (possibly in a nonlinear fashion) [4]. A typical LPV model has the following state space form

$$\dot{x} = A(\theta)x + B(\theta)u \qquad (3.47)$$
$$y = C(\theta)x + D(\theta)u$$

where the time-varying parameter $\theta$ is regarded as an "*exogenous*" real-time measured signal. Typical assumptions on $\theta$ are the bound on magnitude and rate of variation, e.g., $\underline{\theta} \leq \theta \leq \bar{\theta}$ and $\underline{\dot{\theta}} \leq \theta \leq \bar{\dot{\theta}}$ for all $t \geq 0$, with $\underline{\theta}$, $\underline{\dot{\theta}}$ being the lower bounds and $\bar{\theta}$, $\bar{\dot{\theta}}$ the upper bounds. In the subsequent analysis, the notation $\left(\theta, \dot{\theta}\right) \in \Theta \times \Theta$ will be frequently used, where $\Theta$ and $\Theta_d$ are the parameter spaces that contain $\theta$ and $\dot{\theta}$, respectively. Hence, the state space form in (3.4.1) can be regarded as a collection of linear systems (linear differential inclusion).
From a linear perspective, $\theta$ can be considered as a time-varying signal, which is independent of the state variables. Therefore, the state space form given by (3.47) can also be regarded as a linear time-varying (LTV) system. However, the following two points make the LPV system different from the traditional LTV system.

- The parameter $\theta$ can only be measured in real-time as no future parameter values are available. Hence, the control signal is constrained to be a causal function of the parameter. However, some control algorithms for LTV system explicitly use the future information of this parameter [183].
- The LPV framework can be extended to represent a nonlinear dynamical model. In this case, it is called a "quasi-LPV" model, where the time-varying parameter $\theta$ is a nonlinear function of the measurable state variables [183]. In this case, the parameter becomes "*endogenous*". In the LPV model, the complex nonlinearity is hidden behind the time-varying parameter which results in a linear but non-stationary dynamical system. For example, the state equation of an input affine nonlinear system can be transferred to an LPV state equation as

$$\dot{x} = f(x) + g(x)u \Longleftrightarrow \dot{x} = A[\theta(x)]x + B[\theta(x)]u \qquad (3.48)$$

  where $f(x) = A[\theta(x)]$, $g(x) = B[\theta(x)]$ and $\theta(x)$ is a state-dependent measurable parameter.

### Stability analysis of LPV systems

For stability analysis, the following three different methodologies are proposed in the literature to analyze the asymptotic stability of the LPV system [184].

1. Single quadratic Lyapunov function (SQLF) $V = x^T P x$.
2. Parameter dependent quadratic Lyapunov function (PDQLF) $V = x^T P(\theta) x$.
3. Linear fractional representations (LFR) which relies on $\mu$ analysis or small gain theorem for performance optimization and robustness analysis;

Here, an example will be presented to illustrate the application of the 1st approach for stability analysis.
For the LPV system $\dot{x} = A(\theta)x$, where the parameter vector $\theta \in R^{m\times 1}$ belongs to a parameter space $\Theta$. The stability analysis resorts to searching for a Lyapunov function $V = x^T P x$ such that the following LMIs are feasible [185].

$$P \succ 0, \quad A^T(\theta)P + PA(\theta) \prec 0, \quad \forall \theta \in \Theta \qquad (3.49)$$

This is indeed an infinite dimensional LMI feasibility problem due to its continuous dependence on the parameter $\theta$. Fortunately, if the parameter space $\Theta$ is a convex hull with finite vertices and $A(\theta)$ depends affinely on $\theta$, there exists a sufficient and necessary finite dimensional LMI relaxation condition as shown in Lemma 1 [183–185].

**Lemma 3.1**

Suppose the parameter space $\Theta$ is a convex hull with finite vertices and $A(\theta)$ depends affinely on $\theta \in R^{m\times 1}$. Then, the LMIs given by (3.49) are feasible if and only if they are feasible on all the vertices of $\Theta$.

*Proof.* $A(\theta)$ is an affine matrix function of $\theta \in R^{m\times 1}$ as shown below.

$$A(\theta) = A_0 + \sum_{j=1}^{m} \theta_j A_j \tag{3.50}$$

where $\theta_j$ denotes the $j$-th element of $\theta$ . $A_0, A_1, \ldots, A_m$ are all constant matrices. Any parameter $\theta$ in the convex hull $\Theta$ can be represented as a convex combination form as given in (3.51).

$$\theta = \sum_{i=1}^{k} \lambda_i \theta^{(i)} \tag{3.51}$$

where $\theta^{(i)}$ denotes one of the $k$ vertices of $\Theta.\lambda_i, i = 1, \ldots, k$ are the normalized scheduling parameters that satisfy the following convex condition.

$$\sum_{i=1}^{k} \lambda_k = 1, \text{ with } 0 \le \lambda_i \le 1, \forall i = 1 \ldots, k \tag{3.52}$$

Substituting the right side of Eq. (3.51) for $\theta$ in Eq. (3.50), the state matrix $A(\theta)$ can be obtained as the similar convex combination form shown below.

$$\begin{aligned} A(\theta) &= A_0 + \sum_{j=1}^{m} \left( \sum_{i=1}^{k} \lambda_i \theta_j^{(i)} \right) A_j \\ &= A_0 + \sum_{i=1}^{k} \lambda_i \left( \sum_{j=1}^{m} \theta_j^{(i)} A_j \right) \\ &= \sum_{i=1}^{k} \lambda_i A\left(\theta^{(i)}\right) \end{aligned} \tag{3.53}$$

Furthermore, the parameter dependent matrix $A^T(\theta)P + PA(\theta)$ can be represented as

$$A^T(\theta)P + PA(\theta) = \sum_{i=1}^{k} \lambda_i \left[ A^T\left(\theta^{(i)}\right) P + PA\left(\theta^{(i)}\right) \right] \tag{3.54}$$

- **If**: The nonnegativity of $\lambda_i$ in Eq. (3.54) guarantees that $A^T(\theta)P + PA(\theta)$ is negative definite if $A^T(\theta^{(i)})P + PA(\theta^{(i)})$ is a negative definite matrix $\forall i = 1, \ldots, k$.
- **Only if**: The necessity is quite straightforward. Suppose the negative definite condition of $A^T(\theta^{(i)})P + PA(\theta^{(i)})$ is violated at some vertex. Then, $A^T(\theta)P + PA(\theta)$ cannot be a negative definite matrix $\forall \theta \in \Theta$ which leads to a contradiction.

□

### 3.4.2 MHE for nonlinear discrete-time systems

Following the lines of [173], we address the estimation of a discrete-time nonlinear system by adopting the MHE strategy.

#### Problem formulation

Let us consider a dynamic system described by the discrete-time equations

$$x_{k+1} = f(x_k, u_k) + v_k \tag{3.55a}$$

$$y_k = h(x_k) + w_k \tag{3.55b}$$

where $x_k \in \mathbb{R}^n$ is the state vector (the initial state $x_0$ is unknown) and $u_k \in \mathbb{R}^m$ is the control vector, $y_k \in \mathbb{R}^p$ . The vector $w_k \in \mathbb{R}^n$ is an additive disturbance affecting the system dynamics and $v_k$ is the noise in the measurements. We assume the statistics of $x_0$, $v_k$ and $w_k$ to be unknown and consider them as deterministic variables of unknown character that take their values from known compact sets.
The problem consists in estimating, at any time $k = N, N+1, \ldots$, the state vectors $x_{k-N}, \ldots, x_k$ on the basis of a prediction $\bar{x}_{k-N}$ of the state $x_{k-N}$ and of the information vector defined as

$$I_k \triangleq \text{col}\,(y_{t-N}, \ldots, y_t, u_{t-N}, \ldots, u_{t-1}) \tag{3.56}$$

where $N+1$ measurements and $N$ input vectors are collected within a sliding window $[k-N, k]$. Let us define as $\hat{x}_{k-N,k}, \ldots, \hat{x}_{k,t}$ the estimates of $x_{k-N}, \ldots, x_k$, respectively, to be made at time $t$. We assume that the prediction $\bar{x}_{t-N}$ is determined from the estimate $\hat{x}_{k-N-1,k-1}$ via the application of the function f , that is

$$\hat{x}_{i+1,t} = f(\hat{x}_{i,t}, u_i), \quad i = t-N, \ldots, t-1$$

.
The vector $\bar{x}_0$ denotes a priori prediction of $x_0$.
A notable simplification of the estimation scheme can be obtained by defining $\hat{x}_{k-N+1,k}, \ldots, \hat{x}_{k,k}$ as estimates generated by $\hat{x}_{k-N,k}$ through the noise-free dynamics, that is,

$$\hat{x}_{i+1,k} = f(\hat{x}_{i,k}, u_i), \quad i = t-N, \ldots, k-1 \tag{3.57}$$

Hence it follows that at time $k$ only the estimate $\hat{x}_{k-N,k}$ has to be determined, whereas the vectors $\hat{x}_{k-N+1,k}, \ldots, \hat{x}_{k,k}$ can be computed via (3.57).
Let us denote the set of admissible controls by $U$ and the sets from which the vectors $w_k$ and $v_k$ take their values by $\mathcal{W}$ and $\mathcal{V}$, respectively. In order to derive stability results for the estimation error the following assumptions are needed.

**Assumption 3.1**

$\mathcal{W}, \mathcal{V}$, and $\mathcal{U}$ are compact sets, with $0 \in \mathcal{W}$ and $0 \in \mathcal{V}$.

**Assumption 3.2**

The initial state $x_0$ and the control sequence $\{u_k\}$ are such that, for any possible sequence of disturbances $\{w_k\}$, the system trajectory $\{x_k\}$ lies in a compact set $X$.

Let $\mathcal{X}$ be the closed convex hull of the set $X$. The following assumption is also needed.

**Assumption 3.3**

The functions $f$ and $h$ are $\mathcal{C}^2$ functions with respect to $x$ on $\mathcal{X}$ for every $u \in U$.

Note that assumptions 1 and 2 are quite reasonable from a practical point of view when considering state estimation for a physical system: it is very typical that the state variables and the disturbances are bounded in some way. For instance, if assumption 1 is satisfied, then assumption 2 is automatically verified whenever system (3.55a) is input-to-state stable (ISS) with respect to the controls and the disturbances.
We also label as $Y$ and $\mathcal{I}$ the sets from which the vectors $y_k$ and $I_k$ take their values, respectively. Clearly, we have

$$Y = \{y \in \mathbb{R}^p : y = h(x) + v, x \in X, v \in \mathcal{V}\}$$

,

$$\mathcal{I} = Y^{N+1} \times U^N$$

Since, under assumption 2, at every time step $k = 0, 1, \ldots$, the state $x_k$ falls within the set $X$, we consider the condition

$$\hat{x}_{k-N,t} \in X \tag{3.58}$$

The prediction $\bar{x}_{k-N}$ belongs to the compact set $\bar{X} \triangleq f(X, U)$ for every $t = N, N+1, \ldots$ (the a priori prediction $\bar{x}_0$ is chosen inside the set $\bar{X}\Big)$.
The MHE problem consists in deriving the state estimates as functions of the prediction $\bar{x}_{k-N}$ and of the information vector $I_k$. From a formal point of view, this means that we search for a state estimation function of the form $a : \bar{\mathcal{X}} \times \mathcal{I} \mapsto \mathcal{X}$ that provides the estimate
$\hat{x}_{k-N,k} = a\left(\bar{x}_{k-N}, I_k\right)$
Then, the estimation scheme can be summarized as follows

$$\hat{x}_{k-N,k} = a\left(\bar{x}_{k-N}, I_k\right), \quad k = N, N+1, \ldots \tag{3.59a}$$

$$\bar{x}_{t-N+1} = f\left(\hat{x}_{t-N,t}, u_{t-N}\right), \quad t = N, N+1, \ldots \bar{x}_0 \in \bar{X}. \tag{3.59b}$$

As we have assumed the statistics of $x_0, \{\xi_k\}$, and $\{v_k\}$ to be unknown, a natural criterion to derive the state estimation function consists in resorting to a least-squares approach. Toward this end, we consider the following cost function:

$$J\left(\hat{x}_{k-N,k}, \bar{x}_{k-N}, I_k\right) = \mu \left\|\hat{x}_{k-N,k} - \bar{x}_{k-N}\right\|^2 + \sum_{i=k-N}^{k} \left\|y_i - h\left(\hat{x}_{i,k}\right)\right\|^2 \tag{3.60}$$

where $\mu$ is a positive scalar by which we express our belief in the prediction $\bar{x}_{k-N}$ with respect to the observation model. The estimates $\hat{x}_{i,t}, i = t-N+1, \ldots, t$ are obtained as in (3). Then the following estimation algorithm can be stated.
Then, the MHE algorithm is stated as follows:

- **MHE problem**: given an a priori prediction $\bar{x}_0^\circ \in \bar{X}$, at any time step $k = N, N+1, \ldots$, find an estimate $\hat{x}_{k-N,k}^\circ$ such that

$$\hat{x}_{k-N,k}^\circ \in \arg \min_{\hat{x}_{k-N,k} \in X} J\left(\hat{x}_{k-N,k}, \bar{x}_{k-N}^\circ, I_k\right) \tag{3.61}$$

At each step, the prediction is propagated as

$$\bar{x}_{k-N+1}^\circ = f\left(\hat{x}_{k-N,k}^\circ, u_{k-N}\right). \tag{3.62}$$

### Stability of the estimation error

Before stating the results on the stability of the estimation error, the observability properties of the dynamic system (3.55a) observed through the measurement equation (3.55b). In this respect, we shall introduce some useful notations and definitions. Specifically, let us define the function

$$F\left(x_{k-N}, u_{k-N}^{k-1}\right) \triangleq \begin{bmatrix} h\left(x_{k-N}\right) \\ h \circ f^{u_{k-N}}\left(x_{k-N}\right) \\ \vdots \\ h \circ f^{u_{k-1}} \circ \cdots \circ f^{u_{k-N}}\left(x_{k-N}\right) \end{bmatrix}, \tag{3.63}$$

for $k = N, N+1, \ldots$, where $u_{k-N}^{k-1} \triangleq \operatorname{col}\left(u_{k-N}, \ldots, u_{k-1}\right)$, "o" denotes function composition, and $f^{u_i}\left(x_i\right) \triangleq f\left(x_i, u_i\right)$. For the sake of brevity, the vector $u_i$ will be omitted as an argument of the function $f$. Moreover, for the sake of compactness, let $f_{(i)} \triangleq \overbrace{f \circ \cdots \circ f}^{i \text{ times}}$ and

$$f_{(i)}^{\xi_{t-N}^{t-N+i-1}}(x) \triangleq f\left(\cdots f\left(f(x) + \xi_{t-N}\right) + \xi_{t-N+1} \cdots\right) + \xi_{t-N+i-1}.$$

The observability definition is stated below.

**Definition 3.1**

Given a set $S \subseteq \mathbb{R}^n$, system (3.55) is said to be $S$ observable in $N+1$ steps if there exists a $K$-function $\varphi(\cdot)$, such that

$$\varphi\left(\|x_1 - x_2\|^2\right) \leq \|F\left(x_1, \bar{u}\right) - F\left(x_2, \bar{u}\right)\|^2, \forall x_1, x_2 \in S, \quad \forall \bar{u} \in U^N. \tag{3.64}$$

Such a definition has been widely used in the framework of nonlinear state estimation in both the discrete-time and the continuous-time settings (see for example [?, 59, 136] and is related to the properties of the mapping $F(\cdot, \bar{u})$. More specifically, if the algebraic system $F\left(x_{k-N}, u_{k-N}^{t-1}\right) = y_{k-N}^k$ can be uniquely solved by $x_{k-N}$ for any $u_{k-N}^{k-1}$ (i.e., if the mapping $F(\cdot, \bar{u})$ is injective for any $\bar{u} \in U^N$ ), then condition

$$\|F\left(x_1, \bar{u}\right) - F\left(x_2, \bar{u}\right)\|^2 = 0$$

implies $x_1 = x_2$ and hence system (3.55) turns out to be observable in the sense of Definition 1. The following observability assumption is now given.

**Assumption 3.4**

System (3.55) is $\mathcal{X}$-observable in $N+1$ steps with a $K$ function $\varphi$.

Let us now consider a sequence of optimal estimates $\hat{x}^{\circ}_{k-N,k}$, $k = N, N+1, \ldots$ obtained from the MHE algorithm introduced before, and denote the estimation error by $e^{\circ}_{k-N} \triangleq x_{k-N} - \hat{x}^{\circ}_{k-N,k}$. Owing to the definitions and the assumptions given previously, we can state the following results on the stability of the MHE.

**Theorem 3.1: [186]**

Suppose that assumptions 1, 2, 3 and 4 hold. Moreover suppose that the K-function $\varphi$, defined in assumption 4, satisfies the following condition

$$\delta \triangleq \inf_{x_1,x_2\in\mathcal{X};x_1\neq x_2} \frac{\varphi\left(\|x_1 - x_2\|^2\right)}{\|x_1 - x_2\|^2} > 0 \tag{3.65}$$

Then the square norm of the estimation error is bounded as

$$\|e^{\circ}_{k-N}\|^2 \leq \zeta_{k-N} \tag{3.66}$$

where $\{\zeta_k\}$ is a sequence generated by

$$\begin{aligned} \zeta_0 &= \beta_0 \\ \zeta_k &= \alpha\zeta_{k-1} + \beta, \quad t = 1, 2, \ldots \end{aligned} \tag{3.67}$$

with

$$\alpha \triangleq \frac{8k_f^2\mu}{\mu+\delta}$$

,

$$\beta \triangleq \frac{4}{\mu+\delta}\left\{2\mu r_w^2 + \left(\Delta_w\sqrt{N}r_w + \sqrt{N+1}r_v + \frac{\bar{k}}{2}\sqrt{\frac{N(N+1)(2N+1)}{6}r_w^2}\right)^2\right\}$$

$$\beta_0 \triangleq \frac{4}{\mu+\delta}\left\{\mu d_x^2 + \left(\Delta w\sqrt{N}r_w + \sqrt{N+1}r_v + \frac{\bar{k}}{2}\sqrt{\frac{N(N+1)(2N+1)}{6}r_w^2}\right)^2\right\}$$

Moreover, if $\mu$ is selected such that

$$\frac{8k_f^2\mu}{\mu+\delta} < 1, \tag{3.68}$$

the bounding sequence $\{\zeta_k\}$ has the following properties:

i) $\{\zeta_k\}$ converges exponentially to the asymptotic value $e^{\circ}_{\infty}(\mu) \triangleq \beta/(1-\alpha)$

ii) if $\zeta_k > e^{\circ}_{\infty}(\mu)$ then $\zeta_{t+1} < \zeta_t$, $\quad t = 0, 1, \ldots$

*Proof.* For the proof, refer to [186]. □

One may notice that $e^\circ_\infty(\mu)$, which represents an asymptotic upper bound on the quadratic norm of the estimation error, is equal to 0 whenever the system is not affected by system and measurement disturbances($r_w = 0$ and $r_v = 0$). Hence the following corollary stems directly from Theorem 1.

**Corollary 3.1**

Suppose that assumptions 1, 2, 3 and 4 hold. Moreover suppose that the K-function $\varphi$, defined in assumption 4, satisfies condition (3.65). Then, if $w_k = 0$, $v_k = 0$ for $k = 0, 1, \ldots$, the norm of the estimation error is bounded as

$$\left\|e^\circ_{k-N}\right\|^2 \leq \alpha^{k-N}\beta_0,\ \ k = N, N+1, \ldots \tag{3.69}$$

with $\alpha$ and $\beta_0$ defined in Theorem 1. Moreover, if $\mu$ satisfies condition (3.68), then

$$\lim_{k\to\infty} \left\|e^\circ_{k-N}\right\|^2 = 0$$

### 3.4.3 MHE for quasi-LPV systems

Consider the quasi-LPV system described by the following set of equations:

$$x_{k+1} = A\left(p\left(x_k\right)\right)x_k + B\left(p\left(x_k\right)\right)u_k + w_k \tag{3.70a}$$

$$y_k = C\left(p\left(x_k\right)\right)x_k + v_k \tag{3.70b}$$

where $k = 0, 1, \ldots$ is the time instant, $x_k \in \mathbb{R}^n$ is the state vector, $u_k \in \mathbb{R}^q$ is the control vector, $w_k \in \mathbb{R}^n$ is the system noise vector, $y_k \in \mathbb{R}^m$ is the vector of the measures, and $v_k \in \mathbb{R}^m$ is the measurement noise vector. The matrices in (3.70) depend on a set of time-varying parameters that in turn depend on $x_k$ with the mapping $x \mapsto p(x) \in \mathbb{R}^r$ unknown. However, we assume to know the image of such a mapping, namely, the compact set $\mathcal{P} \subset \mathbb{R}^r$ to which the parameters belong (i.e., $p\left(x_k\right) \in \mathcal{P}$ for all $k = 0, 1, \ldots$ ). Thus, in principle, we may rely also on such information for the purpose of estimation.
The systems given by (3.70) include many families of plants and dynamic processes such as those briefly described in the following to highlight the potential advantage as compared with the estimation methods reported in the literature for LPV systems.

A. Uncertain linear systems: linear systems with parameter uncertainties can be viewed as a particular case of (3.70). Such a family of systems is often encountered in the literature because it leads to complicated stabilization problems, as discussed in [187]. The proposed MHE approach enables us to address such issues by estimating simultaneously the uncertain parameters and the state variables.

B. Multi-model systems: the family of multi-model systems under Takagi-Sugeno fuzzy structure [188] is widely investigated in the literature, especially in the research area of fault diagnosis [189]. Such systems can be regarded as belonging to a particular class of plants that can be written under the form (3.70) with bounded unknown parameters. Such unknown parameters, in fuzzy systems, are called premise variables. Generally speaking, when the premise variables are unmeasurable, the estimation and stabilization problems become complicated and from the LMI point of view, the resulting conditions are very conservative. Using the proposed MHE approach, such premise variables can be estimated and then used for the purpose of stabilization.

C. Lipschitz nonlinear systems: one of the major advantages of handling the class of systems given by (3.70) is the fact that such a class includes nonlinear Lipschitz systems often studied in the context of the nonlinear observer design. More specifically, let us consider the Lipschitz nonlinear systems [190] described by

$$\begin{cases} x_{k+1} = \mathcal{A}x_k + B\gamma\left(x_k\right) + w_k \\ y_k = \mathcal{C}x_k + \mathrm{G}\,\sigma\left(x_k\right) + v_k \end{cases} \tag{3.71}$$

where $x_k \in \mathbb{R}^n$ is the state vector; $y_k \in \mathbb{R}^m$ is the output measurement; $w_k \in \mathbb{R}^n$ and $v_k \in \mathbb{R}^m$ are bounded disturbances. The matrices $\mathcal{A} \in \mathbb{R}^{n\times n}, B \in \mathbb{R}^{n\times n_\gamma}, G \in \mathbb{R}^{m\times n_\sigma}$, and $\mathcal{C} \in \mathbb{R}^{m\times n}$ are constant. The function $\gamma : \mathbb{R}^n \longrightarrow \mathbb{R}^{n_\gamma}$ is assumed to be globally Lipschitz and $\gamma(0) = 0$. The same is assumed to hold for $\sigma : \mathbb{R}^n \longrightarrow \mathbb{R}^{n_\sigma}$, i.e., globally Lipschitz with $\sigma(0) = 0$.
Using [ [124], Lemma 2], it follows that there exist functions

$$\phi_{ij} : \mathbb{R}^n \times \mathbb{R}^n \longrightarrow \mathbb{R} \tag{3.72}$$

and constants $a_{ij}, b_{ij}$ such that

$$\gamma\left(x_k\right) = \left(\sum_{i,j=1}^{i,j=n_\gamma,n} \phi_{ij}(k)\mathcal{H}_{ij}\right) x_k \tag{3.73}$$

and

$$a_{ij} \le \phi_{ij}\left(x_k^{0_{j-1}}, x_k^{0_j}\right) \le b_{ij} \tag{3.74}$$

where, for any $X, Y \in \mathbb{R}^n$, we define

$$X^{Y_i} := \left(y_1, \ldots, y_i, x_{i+1}, \ldots, x_n\right) \tag{3.75}$$

for $i = 1, \ldots, n$ with $X^{Y_0} := X$. Similarly, there exist functions

$$\psi_{ij} : \mathbb{R}^n \times \mathbb{R}^n \longrightarrow \mathbb{R} \tag{3.76}$$

and constants $a_{ij}, b_{ij}$, such that

$$\sigma\left(x_k\right) = \left(\sum_{i,j=1}^{i,j=n_\sigma,n} \psi_{ij}(k)\mathcal{H}^\sigma_{ij}\right) x_k \tag{3.77}$$

and

$$c_{ij} \le \psi_{ij}\left(x_k^{0_{j-1}}, x_k^{0_j}\right) \le d_{ij}. \tag{3.78}$$

It follows that (3.71) can be rewritten as an LPV system with

$$\begin{aligned} A\left(p_k\right) &:= \mathcal{A} + B\sum_{i,j=1}^{i,j=n_\gamma,n} \phi_{ij}(k)\mathcal{H}^\gamma_{ij} \\ C\left(p_k\right) &:= \mathcal{C} + G\sum_{i,j=1}^{i,j=n_\gamma,n} \psi_{ij}(k)\mathcal{H}^\sigma_{ij} \end{aligned} \tag{3.79}$$

$p_k := \left(\phi_{11}(k), \ldots, \phi_{1n}(t), \phi_{21}(k), \ldots, \phi_{n_\gamma n}(t), \ldots, \phi_{n_\sigma n}(t)\right)$ where for the sake of shortness,

$$\begin{aligned} \phi_{ij}(k) &:= \phi_{ij}\left(x_k^{0_{j-1}}, x_k^{0_j}\right), \quad \mathcal{H}^\gamma_{ij} := \mathrm{e}_{n_\gamma}(i)\mathrm{e}_n^\top(j), \\ \psi_{ij}(t) &:= \psi_{ij}\left(x_k^{0_{j-1}}, x_k^{0_j}\right), \quad \mathcal{H}^\sigma_{ij} := \mathrm{e}_{n_\sigma}(i)\mathrm{e}_n^\top(j). \end{aligned}$$

**Formulation of the MHE problem**

Two MHE problem can be formulated depending on our trust in the variable parameter $p$. In the first case, a least-squares is considered by minimizing with respect to the parameters which is denoted as the "*optimistic*" approach, and in the second case by maximizing with respect to the parameters which is denoted by the "*pessimistic*" approach. In both cases, we will study the stability of the estimation error, i.e., we will prove the exponential boundedness in the presence of the bounded system and measurement noises.
Let us consider a class of quasi-LPV system described by

$$x_{k+1} = A\left(p\left(x_k\right)\right)x_k + B\left(p\left(x_k\right)\right)u_k + w_k \tag{3.80a}$$

$$y_k = C\left(p\left(x_k\right)\right)x_k + v_k \tag{3.80b}$$

where $k = 0, 1, \ldots$ is the time instant, $x_k \in \mathbb{R}^n$ is the state vector, $w_k \in \mathbb{R}^n$ is the system noise vector, $y_k \in \mathbb{R}^m$ is the vector of the measures, and $v_k \in \mathbb{R}^m$ is the measurement noise vector. $p \in \mathcal{P}$ is the variable parameter, and for the sake of brevity, we denoted $p(x_k)$ by $p_k$. We assume the statistics of $x_0$, $w_k$, and $v_k$ to be unknown, and consider them as deterministic variables of unknown character that take their values from known compact sets.
We can now state the MHE problem according to the two different formulations, i.e., "optimistic" and "pessimistic".

**Problem 1.** Find $\hat{x}^k_{k-N|k} \in \mathbb{R}^{n\times(N+1)}$ and $\hat{p}^k_{k-N|t} \in \mathcal{P}^{N+1}$ that minimize

$$J_1\left(x^k_{k-N}, p^k_{k-N}\right) = \mu\left|x_{k-N} - \bar{x}_{k-N|k}\right|^2 + \sum_{i=k-N}^{k}\left|y_i - C\,x_i\right|^2 \tag{3.81}$$

under the constraints

$$x_{i+1} = A(p_i)\,x_i + B(p_i)\,u_i\,,\; i = k-N, \ldots, k-1\,. \tag{3.82}$$

In practice, the problem consists in finding $\hat{x}^k_{k-N|t}, \hat{p}^k_{k-N|t}$ such that:

$$J_1\left(\hat{x}^k_{k-N|t}, \hat{p}^k_{k-N|k}\right) \leq J_1\left(x^k_{k-N}, p^k_{k-N}\right)$$

for all $x^k_{k-N} \in \mathbb{R}^{n\times(N+1)}$ and $p^k_{k-N} \in \mathcal{P}^{N+1}$.

**Problem 2.** Find $\hat{x}^k_{k-N|k} \in \mathbb{R}^{n\times(N+1)}$ that minimize

$$J_2\left(x^k_{k-N}\right) = \max_{p_{t-N|t}{}^t \in \mathcal{P}^{N+1}} \mu\left|x_{k-N} - \bar{x}_{k-N|k}\right|^2 + \sum_{i=k-N}^{k}\left|y_i - C\,x_i\right|^2 \tag{3.83}$$

under the constraints

$$x_{i+1} = A(p_i)\,x_i + B(p_i)\,u_i\,,\; i = k-N, \ldots, k-1\,. \tag{3.84}$$

### Stability analysis

In order to study the stability properties of the proposed MHE algorithm, some preliminary definitions and assumptions are needed and introduced in the following. Let us define the following function.

$$F\left(p_{k-N}^{k}\right) := \begin{pmatrix} C \\ CA(p_{k-N}) \\ \vdots \\ C\prod_{i=1}^{N} A(p_{k-i}) \end{pmatrix}. \tag{3.85}$$

By exploiting this definition, we get

$$y_{k-N}^{k} = F\left(p_{k-N}^{k}\right)\, x_{k-N} + G\left(p_{k-N}^{t}\right)\, u_{k-N}^{k-1} + H\left(p_{k-N}^{t}\right)\, w_{k-N}^{k-1} + v_{k-N}^{k} \tag{3.86}$$

with $k \geq N$. The matrices $H(.)$ and $G(.)$ are given by (3.87) and (3.88), respectively.

$$H\left(p_{k-N}^{t}\right) := \begin{pmatrix} 0 & \dots & 0 \\ C & \dots & 0 \\ CA(p_{k-N+1}) & \dots & 0 \\ \vdots & \ddots & \vdots \\ C\prod_{i=1}^{N-1} A(p_{k-i}) & \dots & C \end{pmatrix}, \tag{3.87}$$

$$G\left(p_{k-N}^{k}\right) \triangleq H\left(p_{k-N}^{t}\right) \mathrm{diag}\Big(B\left(p_{k-N}\right), \dots, B\left(p_{k}\right)\Big). \tag{3.88}$$

The pair $(A, C)$ is observable in $N$ steps if the following assumption is satisfied

**Assumption 3.5**

the constant

$$\delta := \min_{p_{k-N}^{t} \in \mathcal{P}^{N+1}} \lambda_{\min}\left(F\left(p_{t-N}^{t}\right)^{\top} F\left(p_{k-N}^{t}\right)\right) \tag{3.89}$$

is strictly positive.

In order to derive stability results for the estimation error, we also need that the dynamics of the system to be such that the state vector remains inside some compact set. Therefore, the following assumptions are needed.

**Assumption 3.6**

There exists a compact set $X \subset \mathbb{R}^n$ such that $x_t \in X$ for all $k = 0, 1, \dots,$ and let $\rho_x := \max_{x \in X} |x|$.

**Assumption 3.7**

There exists $\rho_u > 0$ such that $|u_t| \leq \rho_u$ for all $t = 0, 1, \ldots$.

The disturbances are assumed to be bounded as stated in the following assumption.

**Assumption 3.8**

There exist $\rho_w$, $\rho_v > 0$ such that $|w_t| \leq \rho_w$ and $|v_t| \leq \rho_v$ for all $t = 0, 1, \ldots$.

The following assumption is also needed.

**Assumption 3.9**

The mappings $p \mapsto A(p) \in \mathbb{R}^{n\times n}$ and $p \mapsto B(p) \in \mathbb{R}^{n\times m}$ and $p \mapsto C(p) \in \Re^{r\times n}$ are continuous.

Most of the assumptions made so far are of rather technical significance, which will be clear in the proof of the theorem stated in the following.
Before stating the theorem providing a solution to the MHE problem for quasi-LPV systems, we need to introduce the following definition.

**Definition 3.2**

A sequence of vector $\{v_t\}$ is said to be exponentially bounded if there exist $a \in (0, 1)$ and $b > 0$ such that:

$$|v_t| \leq |v_0|a^t + b, \;\; t = 0, 1, \ldots. \tag{3.90}$$

To carry out the stability analysis, we will first provide lower and upper bounds before giving a final bound. Such a final bound will depend on the design parameters, which should be selected conveniently such that the estimation error is exponentially bounded. For the sake of brevity, here we adopt the simpler notations:
$F_{k-N} := F\left(p^t_{k-N}\right), G_{k-N} := G\left(p^k_{k-N}\right)$, $H_{k-N} := H\left(p^k_{t-N}\right)$, $\hat{F}_{k-N} := F\left(\hat{p}^k_{k-N}\right)$, $\hat{G}_{k-N} := G\left(\hat{p}^k_{k-N}\right)$, $\hat{H}_{k-N} := H\left(\hat{p}^k_{k-N}\right)$, and $\hat{x}_i := \hat{x}_{i|k}$, $\hat{p}_i := \hat{p}_{i|k}$, for $i = k - N, \ldots, k$.
Let us consider the optimal cost that results from the solution of problem 1 defined before, namely,

$$\begin{aligned} J_1\left(\hat{x}^k_{k-N}, \hat{p}^k_{k-N}\right) = \mu \left|\hat{x}_{k-N} - \bar{x}_{k-N|k}\right|^2 + \Big| y^k_{k-N} - \hat{F}_{k-N}\,\hat{x}_{k-N} \\ -\hat{G}_{k-N-1}u^{t-1}_{k-N-1}\Big|^2 . \end{aligned} \tag{3.91}$$

Likewise in [138, 174], the proof consists in deriving lower and upper bounds on such a cost, which will be combined define the final bound on the norm of the estimation error.

- **Lower bound**

First, we bound the first term of (3.91). We have:

$$\begin{aligned}|x_{k-N} - \hat{x}_{k-N}| &= |x_{k-N} - \bar{x}_{k-N|t} + \bar{x}_{k-N|k} - \hat{x}_{k-N}| \\ &\leq |x_{k-N} - \bar{x}_{k-N|t}| + |\bar{x}_{k-N|t} - \hat{x}_{k-N}|.\end{aligned}$$

Using Young's inequality we get:

$$|x_{k-N} - \hat{x}_{k-N}|^2 \leq (1 + \frac{1}{\epsilon_1})|x_{k-N} - \bar{x}_{k-N|k}|^2 + (1 + \epsilon_1)|\bar{x}_{k-N|k} - \hat{x}_{t-N}|^2$$

and therefore

$$\mu|\bar{x}_{k-N|k} - \hat{x}_{k-N}|^2 \geq \frac{\mu}{1+\epsilon_1}|x_{-N} - \hat{x}_{-N}|^2 - \frac{\mu}{\epsilon_1}|x_{-N} - \bar{x}_{-N|}|^2 \tag{3.92}$$

for any $\mu \geq 0$ and $\epsilon_1 > 0$.

Now, we focus on the second term of (3.91):

$$\begin{aligned}\left|\hat{F}_{k-N}\, x_{k-N} - \hat{F}_{k-N}\, \hat{x}_{k-N}\right| &= \left|\hat{F}_{k-N}\, x_{k-N} - F_{k-N}\, x_{k-N} + F_{k-N}\, x_{k-N} - y_{k-N+1}^{k} + \hat{G}_{k-N} u_{k-N}^{k-1}\right. \\ &\quad \left. + y_{k-N+1}^{k} - \hat{F}_{k-N}\, \hat{x}_{k-N} - \hat{G}_{k-N} u_{k-N}^{k-1}\right| \\ &\leq \left|\left(\hat{F}_{k-N} - F_{k-N}\right)\, x_{k-N}\right|\left|F_{k-N}\, x_{k-N} - y_{k-N+1}^{k} + \hat{G}_{k-N} u_{k-N}^{k-1}\right| \\ &\quad + \left|y_{k-N+1}^{k} - \hat{F}_{k-N}\, \hat{x}_{k-N} - \hat{G}_{k-N} u_{k-N}^{k-1}\right|. \\ &\leq \left|y_{k-N+1}^{k} - \hat{F}_{k-N}\, \hat{x}_{k-N} - \hat{G}_{k-N} u_{k-N}^{k-1}\right| \\ &\quad + \underbrace{\Delta_F \rho_x + \rho_H N \rho_w + (N+1)\rho_v + \rho_u N \Delta_G}_{c_1}\end{aligned} \tag{3.93}$$

where

$$\begin{aligned}\Delta_F &:= \max_{p1_{k-N}^{k}, p2_{k-N}^{k} \in \mathcal{P}^{N+1}} \left|F\left(p1_{k-N}^{k}\right) - F\left(p2_{k-N}^{k}\right)\right| \\ \Delta_G &:= \max_{p1_{k-N}^{k}, p2_{k-N}^{k} \in \mathcal{P}^{N+1}} \left|G\left(p1_{k-N}^{k}\right) - G\left(p2_{k-N}^{k}\right)\right| \\ \rho_H &:= \max_{p_{k-N}^{k} \in \mathcal{P}^{N+1}} \left|H\left(p_{k-N}^{k}\right)\right|,\ \rho_u := \max_k |u_k|.\end{aligned}$$

From Young's inequality we obtain:

$$\begin{aligned}\left|\hat{F}_{k-N\ k-N} - \hat{F}_{k-N}\, \hat{x}_{k-N}\right|^2 &\leq (1+\epsilon_2)\left|y_{k-N+1}^{k} - \hat{F}_{k-N}\, \hat{p}_{k-N} - \hat{G}_{k-N} u_{k-N}^{k-1}\right|^2 \\ &\quad + c_1^2(1 + \frac{1}{\epsilon_2}).\end{aligned} \tag{3.94}$$

Using Assumption 5, we obtain

$$\begin{aligned}\left|y_{k-N+1}^{k} - \hat{F}_{k-N}\, \hat{x}_{k-N} - \hat{G}_{k-N} u_{k-N}^{k-1}\right|^2 &\geq \frac{1}{(1+\epsilon_2)}\left|\hat{F}_{k-N}\, \left(x_{k-N} - \hat{x}_{k-N}\right)\right|^2 - \frac{1}{\epsilon_2}c_1^2 \\ &\geq \frac{\delta}{1+\epsilon_2}|x_{k-N} - \hat{x}_{k-N}|^2 - \frac{1}{\epsilon_2}c_1^2.\end{aligned} \tag{3.95}$$

To conclude, by replacing (3.92) and (3.95) in (3.91), we get

$$J_1\left(\hat{x}^k_{k-N}, \hat{p}^k_{k-N}\right) \geq \frac{\mu}{1+\epsilon_1}|x_{k-N} - \hat{x}_{k-N}|^2 - \frac{\mu}{\epsilon_1}|x_{k-N} - \bar{x}_{k-N|k}|^2 + \frac{\delta}{1+\epsilon_2}|x_{k-N} - \hat{x}_{k-N}|^2 - \frac{c_1^2}{\epsilon_2}\,. \tag{3.96}$$

- **Upper Bound**

It is straightforward to proceed with the upper bound as follows:

$$\begin{aligned} J\left(\hat{q}^k_{k-N}, \hat{p}^k_{k-N}\right) &\leq J\left(q^k_{k-N}, p^k_{k-N}\right) \\ &= \left|y^k_{k-N+1} - F_{k-N}\, q_{k-N} - G_{k-N} u^{k-1}_{k-N}\right|^2 + \mu\, |q_{k-N} - \bar{q}_{k-N|k}|^2 \\ &= \mu\, |q_{k-N} - \bar{q}_{k-N|k}|^2 + \left|H_{k-N}\, w^{k-1}_{k-N} + v^k_{k-N}\right|^2 \\ &\leq \mu\, |q_{k-N} - \bar{q}_{k-N|k}|^2 + 2\left|H_{k-N}\, w^{k-1}_{k-N}\right|^2 + 2\left|v^k_{k-N}\right|^2 \\ &\leq \mu\, |q_{k-N} - \bar{q}_{k-N|k}|^2 + c_2, \end{aligned} \tag{3.97}$$

where

$$c_2 := 2\,\rho_H^2 N^2 \rho_w^2 + 2(N+1)^2 \rho_v^2\,.$$

- **Final bound**

Using (3.96) and (3.97), we get:

$$\left(\frac{\delta}{1+\epsilon_2} + \frac{\mu}{1+\epsilon_1}\right)|q_{k-N} - \hat{q}_{k-N}|^2 \leq \frac{c_1^2}{\epsilon_2} + c_2 + \mu(1 + \frac{1}{\epsilon_1})\,|q_{k-N} - \bar{q}_{k-N|k}|^2. \tag{3.98}$$

Let us now focus on the term $|q_{k-N} - \bar{q}_{k-N}|$ in the r.h.s. of (3.105). It follows that:

$$\begin{aligned} |q_{k-N} - \bar{q}_{k-N}| &= \Big|A(p_{k-N-1})\, q_{k-N-1} + w_{k-N-1} - A(\hat{p}_{k-N-1})\, \hat{q}_{k-N-1} \\ &\quad + \left(B(p_{k-N-1}) - B(\hat{p}_{k-N-1})\right) u_{k-N-1}\Big| \\ &= \Big|\left(A(p_{k-N-1}) - A(\hat{p}_{k-N-1})\right) q_{k-N-1} + A(\hat{p}_{k-N-1})\left(q_{k-N-1} - \hat{q}_{k-N-1}\right) \\ &\quad + \left(B(p_{k-N-1}) - B(\hat{p}_{k-N-1})\right) u_{k-N-1} + w_{k-N-1}\Big| \\ &\leq |\left(A(p_{k-N-1}) - A(\hat{p}_{k-N-1})\right) q_{k-N-1}||A(\hat{p}_{k-N-1})\left(q_{k-N-1} - \hat{q}_{k-N-1}\right)| \\ &\quad + |\left(B(p_{k-N-1}) - B(\hat{p}_{k-N-1})\right) u_{k-N-1}| + |w_{k-N-1}|. \end{aligned}$$

Using Young's inequality, we get

$$|q_{k-N} - \bar{q}_{k-N}|^2 \leq \frac{1+\epsilon_3}{\epsilon_3}\left(\Delta_A \rho_q + \rho_w + \Delta_B \rho_u\right)^2 + (1+\epsilon_3)\rho_A^2 |q_{k-N-1} - \hat{q}_{k-N-1}|^2 \tag{3.99}$$

for all $\epsilon_3 > 0$, where

$$\Delta_A := \max_{p_1, p_2 \in \mathcal{P}} |A\left(p_1\right) - A\left(p_2\right)|\,, \tag{3.100}$$

$$\Delta_B := \max_{p_1, p_2 \in \mathcal{P}} |B\left(p_1\right) - B\left(p_2\right)|\,. \tag{3.101}$$

By substituting (3.106) in (3.105), we finally obtain

$$|\tilde{x}_{k-N}|^2 \leq \left( \frac{\mu(1+\frac{1}{\epsilon_1})(1+\epsilon_3)}{\frac{\delta}{1+\epsilon_2}+\frac{\mu}{1+\epsilon_1}} \right) \rho_A^2 |\tilde{x}_{k-N-1}|^2 + \left( \frac{\frac{c_1^2}{\epsilon_2}+c_2+\frac{1+\epsilon_3}{\epsilon_3}\mu\sigma^2}{\frac{\delta}{1+\epsilon_2}+\frac{\mu}{1+\epsilon_1}} \right) \tag{3.102}$$

where $\sigma \triangleq \Delta_A \rho_x + \rho_w + \Delta_B \rho_u$ and $\tilde{x}_{k-N} \triangleq x_{k-N} - \hat{x}_{k-N}$ .
Now we are ready to state the following main theorem, which summarizes the convergence conditions of the estimation error obtained from the solution of the optimization problem 1.

**Theorem 3.2**

The estimation error $e_{k-N}$ given by the solution of MHE problem is exponentially bounded according to Definition 2, with

$$a = a(\mu) \triangleq \rho_A \sqrt{\frac{\mu(1+\frac{1}{\epsilon_1})(1+\epsilon_3)}{\frac{\delta}{1+\epsilon_2}+\frac{\mu}{1+\epsilon_1}}} \tag{3.103}$$

$$b = b(\mu) \triangleq \left( \frac{\frac{c_1^2}{\epsilon_2}+c_2+\frac{1+\epsilon_3}{\epsilon_3}\mu\sigma^2}{\frac{\delta}{1+\epsilon_2}+\frac{\mu}{1+\epsilon_1}} \right) \frac{1}{1-a} \tag{3.104}$$

if $\mu$ is chosen such that $a(\mu) < 1$.

*Proof.* Using (3.96) and (3.97), we get

$$\left( \frac{\delta}{1+\epsilon_2} + \frac{\mu}{1+\epsilon_1} \right) |x_{t-N} - \hat{x}_{t-N}|^2 \leq \frac{c_1^2}{\epsilon_2} + c_2 + \mu(1 + \frac{1}{\epsilon_1}) |x_{k-N} - \bar{x}_{k-N|k}|^2. \tag{3.105}$$

Let us now focus on the term $|x_{t-N} - \bar{x}_{k-N}|$ in the r.h.s. of (3.105). It follows that:

$$\begin{aligned} |x_{k-N} - \bar{x}_{k-N}| &= \Big| A(p_{k-N-1})\, x_{k-N-1} + w_{k-N-1} - A(\hat{p}_{k-N-1})\, \hat{x}_{k-N-1} \\ &\quad + (B(p_{k-N-1}) - B(\hat{p}_{k-N-1}))\, u_{k-N-1} \Big| \\ &= \Big| (A(p_{k-N-1}) - A(\hat{p}_{k-N-1}))\, x_{k-N-1} + A(\hat{p}_{k-N-1})\, (x_{k-N-1} - \hat{x}_{k-N-1}) \\ &\quad + (B(p_{k-N-1}) - B(\hat{p}_{k-N-1}))\, u_{k-N-1} + w_{k-N-1} \Big| \\ &\leq |(A(p_{k-N-1}) - A(\hat{p}_{k-N-1}))\, x_{k-N-1}| + |A(\hat{p}_{k-N-1})\, (x_{k-N-1} - \hat{x}_{k-N-1})| \\ &\quad + |(B(p_{k-N-1}) - B(\hat{p}_{k-N-1}))\, u_{k-N-1}| + |w_{k-N-1}|. \end{aligned}$$

Using Young's inequality, we get

$$|x_{k-N} - \bar{x}_{k-N}|^2 \leq \frac{1+\epsilon_3}{\epsilon_3} \Big( \Delta_A \rho_x + \rho_w + \Delta_B \rho_u \Big)^2 + (1+\epsilon_3)\rho_A^2 |x_{k-N-1} - \hat{x}_{k-N-1}|^2 \tag{3.106}$$

for all $\epsilon_3 > 0$, where

$$\Delta_A := \max_{p_1, p_2 \in \mathcal{P}} |A(p_1) - A(p_2)| \ . \tag{3.107}$$

$$\Delta_B := \max_{p_1,p_2\in\mathcal{P}} |B(p_1) - B(p_2)| \, . \tag{3.108}$$

By substituting (3.106) in (3.105), we finally obtain:

$$|\tilde{x}_{k-N}|^2 \le \left(\frac{\mu(1+\frac{1}{\epsilon_1})(1+\epsilon_3)}{\frac{\delta}{1+\epsilon_2}+\frac{\mu}{1+\epsilon_1}}\right)\rho_A^2|\tilde{x}_{k-N-1}|^2 + \left(\frac{\frac{c_1^2}{\epsilon_2}+c_2+\frac{1+\epsilon_3}{\epsilon_3}\mu\sigma^2}{\frac{\delta}{1+\epsilon_2}+\frac{\mu}{1+\epsilon_1}}\right) \tag{3.109}$$

where $\sigma \triangleq \Delta_A\rho_x + \rho_w + \Delta_B\rho_u$ and $\tilde{x}_{k-N} \triangleq x_{k-N} - \hat{x}_{k-N}$ . □

**Remark 3.4**

If we take $\epsilon_i = 1, i = 1, 2, 3$, then estimation error bound are simplified as follows:

$$a(\mu) = \frac{8\mu}{\delta+\mu}\rho_A^2 \tag{3.110}$$

$$b(\mu) = \frac{2c_1^2 + 2c_2 + 4\mu\sigma^2}{(\delta+\mu)(1-a)}. \tag{3.111}$$

Now let us focus on the stability analysis of the estimation error given by the solution of Problem 2.

**Theorem 3.3**

The estimation error $e_{k-N}$ given by the solution of Problem 2 is exponentially bounded with the same definitions of $a_2(\mu)$ in (3.103) and $b_2(\mu)$ in (3.104) if $\mu$ is chosen such that $a_2(\mu) < 1$.

*Proof.* The proof is straightforward, we follow the same procedures as in the proof of Theorem 2 using the cost function

$$J_2\left(\hat{x}^k_{k-N|k}\right) = J_1\left(\hat{x}^k_{k-N|k}, \hat{\pi}^k_{k-N|k}\right)$$

where $\hat{\pi}^k_{k-N|k} \in \mathcal{P}$ is the maximizer of the right half side of (3.83) with the corresponding definitions ($F_{k-N} = F\left(\pi^k_{k-N}\right)$, $H_{k-N} = H\left(\pi^k_{k-N}\right)$). □

**Remark 3.5**

It is important to note that the MHE is able to provide estimates of both system states and unknown parameters, even when such parameters affect the system matrices nonlinearly. Indeed, the matrices $A$, $B$ and $C$ in (3.70) may depend nonlinearly on the parameter $p_k$. Although owing to the boundedness of the parameters, it is always possible to avoid such nonlinearities by introducing a new extended parameter vector, $\bar{p}_k$, with higher dimension and rewrite $A$, $B$ and $C$ with a linear dependence on $\bar{p}_k$, this would increase the size of the new parameter $\bar{p}_k$. If such a parameter increases significantly, it may happen also to lose the detectability and stabilizability conditions when augmenting the size of the unknown parameter vector as well as the infeasibility of LMI conditions ensuring the stability of the estimation error for other alternative estimation methods for LPV systems. Moreover, LMI-based techniques may require additional conservative conditions to guarantee the existence of an observer. This is the case of adaptive observers with strong equality constraints [191] or unknown input observers with restrictive rank conditions based on singular systems theory [192].

## 3.5 Conclusion

In this chapter, a brief overview of the optimization-based strategies for estimating the state of linear and nonlinear systems has been reviewed. These methods can provide optimal estimators for linear systems and stable estimators for constrained linear and finite-time observable nonlinear systems. The use of moving horizon estimation (MHE) for state estimation is then presented which provides an estimate at the current instant by solving an optimization problem based on information from a fixed number of the latest measurements collected over a finite horizon. In the second part of this chapter, the MHE was investigated for a class of nonlinear systems that can be written under the form of quasi-LPV system to estimate jointly the system's states and the unknown parameters. The MHE is accomplished by minimizing a least-squares cost function with respect to both state and parameter variables. In the first approach, the unknown parameters are regarded as state variables, hence minimizing with respect to both state and parameter variables which is referred as the "optimistic" approach. In the second approach, the unknown parameters are considered in the worst case, by solving a min-max problem referred as the "pessimistic" approach. Sufficient conditions are drawn to guarantee the stability of the estimation error.

Chapter 4

# Applications and simulation

*"In theory, theory and practice are the same. In practice, they are not."*

Albert Einstein]

## Contents



## 4.1 Introduction

Understanding biological systems is difficult as they are susceptible to adapt (their properties evolve over time), or can be highly sensitive to a small change in their environment, it is, therefore, crucial to monitor such adapting biological processes. However, these biological models usually introduce some unknown parameters that must be estimated accurately and efficiently. Hence, one central challenge in systems biology is the estimation of unknown model parameters (e.g., rate constants), after which one can predict model dynamics. Parameter estimation requires the knowledge of the system's state variables, but due to technical limitations, only part of the state variables can be measured in experiments due to a lack of sensors sometimes, these sensors are expensive and requires high expertise and strong maintenance. Moreover, most of the measurements are generally obtained offline, after laboratory analysis.

As most models for molecular biological systems are nonlinear with respect to both parameters and system state variables, the estimation of parameters in these models from experimental measurement data is thus a nonlinear estimation problem. In principle, all algorithms for nonlinear optimization can be used to deal with this problem, for example, the Gauss-Newton iteration method and its variants. However, these methods do not take the special structures of biological system models into account. When the number of states or parameters to be determined increases, it will be challenging and computationally expensive to apply these conventional methods. One way to obtain such variables consists of combining a priori knowledge about these biological systems with experimental data to design observers which process the incomplete and imperfect information to construct an estimate of these unknown parameters and state variables.

In this chapter, the high-gain estimation techniques introduced in Chapter 2 and the moving horizon estimation approach presented in Chapter 3 are investigated to estimate the unknown state variables and parameters of three different biological plants. First, the applicability of the high-gain observers based on the system state augmentation approach and HG/LMI technique is verified on two biological applications, namely, a simple one-gene regulation dynamic process involving end-product activation is explored to estimate the non-measured concentrations of mRNA, small metabolites, and the involved protein, then a stage-structured SI epidemic model with 2 infectious stages is considered to track the states of the epidemic model. Simulation results are reported showing the superiority of our proposed observer in terms of sensitivity to high-frequency measurement noise with respect to the standard high-gain observer in addition to improving the transient response by preventing the peaking phenomenon which is typical of the high-gain observer. Second, we have investigated both modeling and estimation of the planar dynamics of Amnioserosa cells during dorsal closure. The model is described as a quasi-LPV system, and the moving horizon estimator proposed in Chapter 3 is considered to estimate the system's states and the unknown parameters jointly. For simulation purposes, data from the work published in [193] are used to evaluate the performance of the MHE which is compared to the traditional extended Kalman filter (EKF). The comparative analysis revealed that not only the MHE technique can achieve optimal state and parameter-estimation performances comparable to the EKF, but also more robust to measurement and process noises.

## 4.2 Genetic regulatory network

Gene expression is a very complicated dynamical process that is regulated at a number of stages during the synthesis of proteins [194]. Similar to many big cities, with heavy traffic, biological cells host complicated traffic of biochemical signals at all levels. At the nanometer scale, 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 [195]. During gene expression, an enormous number of genes and proteins are either directly or indirectly interrelated with one another in living cells. The mechanisms which have progressed to control the expression of genes are known as gene regulatory networks (GRNs). Gene Regulatory Networks (GRNs) describe gene expression as a function of regulatory inputs specified by the interactions between proteins and genes, that is, one gene can control another gene's expression through its product proteins called Transcriptional Factors [196]. Understanding GRNs has the potential to advance fields ranging from basic science research to clinical practice and provides insights in evolution [197], metabolism [198], DNA damage response [199], and cancer metastasis [200].

For the purpose of gene identification and medical diagnosis/treatment, biologists and biomedical scientists are interested in knowing the exact states of GRNs [201]. However, due to unavoidable complications such as transcription/translation delays and extrinsic/intrinsic noises, the available measurement outputs of GRNs might be different from their true states. As such, the state estimation or filtering problem for GRNs has become an important topic of research that has attracted many from the scientific community [202], [203], [201], [204], [205] and [206].

In this section, The applicability and performance of the high-gain observer techniques addressed in the Chapter 2 are explored through a simple genetic regulatory network with a comparison to the standard high-gain observer to clearly show the superiority of the proposed technique.

### 4.2.1 Mathematical modeling of genetic regulation process

Mathematical modeling has been applied to biological systems for decades, but with respect to gene expression, too few molecular components have been known to build useful, predictive models. New efforts have been greatly aided by much more extensive "parts lists" of DNA sequences and proteins, as well as considerably enhanced computational power. These improvements make possible the use of diverse mathematical modeling methods for different biological problems. As more biologists venture into systems-level studies, a general understanding of various modeling approaches related to gene expression is necessary to facilitate close collaborations between experimentalists and modelers [207]. There are different classes of genetic regulatory models, namely, thermodynamic, Boolean, Bayesian, and differential equation-based models. In differential equation based-models, the regulatory networks can be represented by differential equations, in which a set of molecules such as mRNAs and proteins interact by explicit rules defined in terms of rate equations. These equations specify the levels of each protein or mRNA as a function of the other components as the system evolves. These models usually include time and/or space-dependent variables such as protein and mRNA concentrations, and parameters such as production and degradation rates. A Regulatory relationship between two genes is depicted in Figure 4.1, where synthesis of gene 1 ($G_1$) involves expression of mRNA ($M_1$) and translation of protein ($P_1$), which regulates gene 2 ($G_2$). Both mRNA and protein are subject to turnover and protein are subject to diffusion, mRNA and protein synthesis, degradation, and diffusion events are shown at left. This process can be modeled with reaction-diffusion equations given in the figure and each molecular constituent is assigned such an equation.

The ordinary differential equation (ODE) formalism models the concentrations of mRNAs, proteins, and other system elements by time-dependent variables with values contained in the set of non-negative real numbers. Regulatory interactions take the form of functional and differential relations between the concentration variables. More specifically, gene regulation is modeled by reaction-rate equations expressing the rate of elements of the system in the following form:

$$\frac{dx_i}{dt} = f_i(x),\ x_i \geq 0,\ 1 \leq i \leq n, \tag{4.1}$$

where $x$ is the vector of concentrations of proteins, mRNAs, or small metabolites, and $f_i$ is a usually nonlinear function. Then, by specifying the function $f$ we get equations of the form 4.2 [208]:

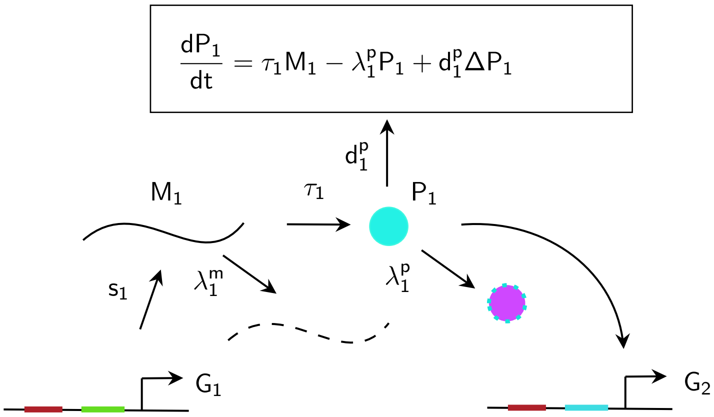



Figure 4.1: Differential equation model of gene expression.

$$\begin{cases} \frac{dx_1}{dt} = \kappa_{1n} H(x_n) - \gamma_1 x_1, \ x_1 \geq 0, \\ \frac{dx_i}{dt} = \kappa_{i,i-1} x_{i-1} - \gamma_i x_i, \ x_i \geq 0, \ 1 \leq i \leq n \end{cases} \tag{4.2}$$

The parameters $\kappa_{1n}, \kappa_{2n}, \cdots \kappa_{n,n-1}$ are all strictly positive and represent production constants and $\gamma_1, \cdots \gamma_n$ represent degradation constants and are also strictly positive. The reaction-rate equations express a balance between the number of molecules appearing and disappearing per unit of time. In the case of $x_1$, the production term involves a nonlinear regulation function $H : \mathbb{R} \to \mathbb{R}$ ranging from 0 to 1. A regulation function often found in the literature is the so-called Hill curve depicted in Figure 4.2 with $m > 0$ as the steepness parameter.

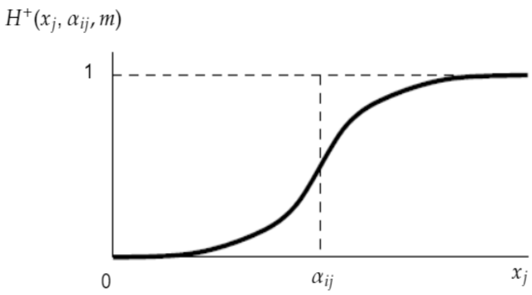



Figure 4.2: Hill function $H^+$.

In common use, the function $H$ is given as

$$H^+(x_j, \theta_{ij}, m) = \frac{x_j^m}{x_j^m + \alpha_{ij}^m},$$
$$H^-(x_j, \theta_{ij}, m) = 1 - \frac{x_j^m}{x_j^m + \alpha_{ij}^m},$$

corresponding to the activation and inhibition cases, respectively. The parameter $\alpha$ gives the threshold for the regulatory influence of the concentration of the metabolite on the target gene, whereas the steepness parameter $m$ is a measure of the collective effect of groups of metabolite molecules defining the shape of the Hill curve.

### 4.2.2 Simple gene regulation process

A simple example of a kinetic model for the genetic regulation process is given in Figure 4.3, going back to the work by Goodwin [209] and [210]. The end-product of a metabolic pathway co-inhibits the expression of a gene coding for an enzyme that catalyzes a reaction step in the pathway. This gives rise to a negative feedback loop involving mRNA concentration $x_1$, protein concentration $x_2$, and metabolic concentration $x_3$.

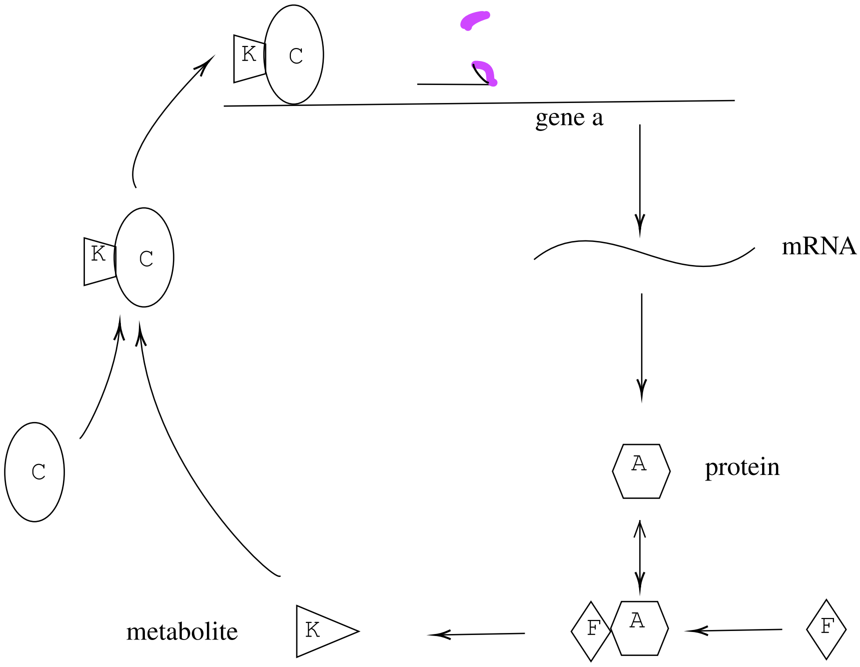



Figure 4.3: An example of a genetic regulatory system involving end-product inhibition. A and C represent proteins, F and K metabolites.

Let $x_1$, $x_2$ and $x_3$ be the concentrations of the messenger RNA (mRNA) a, protein A and metabolite K, respectively. Then, the corresponding Tysonâ€™s model [208] is written under the form:

$$\Gamma_{GRN} : \begin{cases} \dot{x}_1 = \kappa_1 H(x_3) - \gamma_1 x_1 \\ \dot{x}_2 = \kappa_2 x_1 - \gamma_2 x_2 \\ \dot{x}_3 = \kappa_3 x_2 - \gamma_3 x_3 \\ y = x_3 \end{cases} \tag{4.3}$$

$\kappa_1$, $\kappa_2$ and $\kappa_3$ are production constants, $\gamma_1$, $\gamma_2$ and $\gamma_3$ degradation constants, and $H$ is the nonlinear regulation function ranging from 0 to 1.
System (4.3) is considered to be a good model for the simplest type of allosteric regulation in biochemistry, i.e., the inhibition or activation of an enzyme or protein by a small regulatory molecule that interacts with the enzyme at a site (allosteric site) other than the active site at

Table 4.1: Parameters of the GRN

| Symbol | Meaning | Value(arb. units) |
|---|---|---|
| $\kappa_1$ | Production constant of mRNA | 0.001 |
| $\kappa_2$ | Production constant of protein | 1.0 |
| $\kappa_3$ | Production constant of metabolite | 1.0 |
| $\gamma_1$ | Degradation constant of mRNA | 0.1 |
| $\gamma_2$ | Degradation constant of protein | 1.0 |
| $\gamma_3$ | Degradation constant of metabolite | 1.0 |
| $\alpha$ | Hill's threshold parameter | 1.0 |

which catalytic activity occurs. The interaction changes the shape of the enzyme, thus affecting the active site of the standard catalysis. This change of shape of the enzyme is sufficient to change its ability to catalyze a reaction in either a negative or positive way and enables a cell to regulate needed metabolites. The allosteric regulation has the typical features of a feedback loop in control theory if the regulatory protein acts on the enzyme in the pathway of its own synthesis.
The model (4.3) has the following form:

$$\begin{cases} \dot{x} = f(x) \\ y = h(x) \end{cases} . \tag{4.4}$$

Using an appropriate change of coordinate, the system is transformed into a triangular form as given in (4.5):

$$\begin{cases} \dot{\xi} = F'(\xi) = \begin{bmatrix} \dot{\xi}_1 \\ \dot{\xi}_2 \\ \dot{\xi}_3 \end{bmatrix} = \begin{bmatrix} \xi_2 \\ \xi_3 \\ \varphi(\xi) \end{bmatrix}, \\ y = C\xi = \begin{bmatrix} 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} \xi_1 \\ \xi_2 \\ \xi_3 \end{bmatrix} \end{cases} \tag{4.5}$$

The standard high-gain observer and the observer based on a combination of HG/LMI technique and system state augmentation approach presented in Chapter 2 are designed to estimate the states of the system given by (4.5). For $j_s = 1$, we transform the system given by (4.5) of dimension 3 to an augmented system of dimension 4 by adding an integrator to obtain this fourth-dimensional system:

$$\begin{cases} \dot{z}_1 = z_2 \\ \dot{z}_2 = z_3 \\ \dot{z}_3 = z_4 \\ \dot{z}_4 = \varphi'(\xi) \\ y = z_1 \end{cases} \tag{4.6}$$

where $\varphi'(.)$ is the new nonlinear function that does depend on the last component $z_4$.

### 4.2.3 Simulation results

In order to carry out simulations, we have used the parameters given in Table 4.1 which are not necessarily the experimental values but are consistent with the requirements of the model.
Following theorem 4 we have designed a standard high-gain observer for system4.5, then using theorem 8 we have designed the observer proposed in this work for system 4.5. Note that in

Table 4.2, the line of the first sub-table "Improved high-gain observer" with $j_s = 0$, $j_s = 0$ refers to the worst case of our proposed observer which corresponds to the standard high-gain observer. The two remaining lines correspond to the HG/LMI observer with a compromise index $j_0 = 1$ and $j_0 = 2$. For the second sub-table, we have the case of the augmented system to a dimension 4 $(3 + j_s = 4)$ combined with the HG/LMI observer for three different values of the compromise index $j_0 = 0, 1, 2$.

Table 4.2: Comparison between the two high-gain observers for the GRN model.

| **Standard high-gain observer** | | | | |
|---|---|---|---|---|
| $n_{\text{LMI}}$ | $\theta$ | $K$ | | |
| 1 | 24.23 | 8.87 | 15.79 | 7.91 |

| **Improved high-gain observer** | | | | | | |
|---|---|---|---|---|---|---|
| $j_s$ | $j_0$ | $n_{\text{LMI}}$ | $\theta_\Gamma^{\frac{1}{(1+j_s)(1+j_0)}}$ | $K_\Gamma$ | | |
| 0 | 0 | 1 | 24.23 | 8.87 | 15.79 | 7.91 |
| | 1 | 2 | 4.61 | 8.72 | 15.99 | 8.28 |
| | 2 | 4 | 2.71 | 7.27 | 14.77 | 7.22 |

| $j_s$ | $j_0$ | $n_{\text{LMI}}$ | $\theta_\Gamma^{\frac{1}{(1+j_s)(1+j_0)}}$ | $K_\Gamma$ | | | |
|---|---|---|---|---|---|---|---|
| 1 | 0 | 1 | 7.55 | 8.72 | 18.59 | 18.06 | 7.27 |
| | 1 | 2 | 2.53 | 8.94 | 19.27 | 18.91 | 7.81 |
| | 2 | 4 | 1.89 | 9.18 | 20.02 | 19.93 | 8.43 |

Figure 4.4, depicts the behavior of the error dynamics $x_i - \hat{x}_i$, $i = 1, \ldots, 3$ using the standard high-gain observer and our proposed observer obtained from the combination of the state augmentation approach and the HG/LMI technique. Concerning our proposed observer, we have picked three case studies:

- case 1 ($j_s = 0$, $j_0 = 1$): corresponding to the HG/LMI observer with the compromise index $j_0 = 1$ but without augmenting the state of the original system.
- case 2 ($j_s = 1$, $j_0 = 0$): corresponding to the observer based on the system state augmentation approach with the compromise index $j_s = 1$ but without applying the HG/LMI technique on the augmented system.
- case 3 ($j_s = 1$, $j_0 = 2$): which corresponds to the case where we increase the dimension of the system by 1 by augmenting the state combined with the HG/LMI technique with the compromise index $j_0 = 2$.

We denote $\hat{x}_{i,s}$, the system's estimate using the standard high-gain observer which is depicted by a solid blue line, $\hat{x}_{i,e_{(j_s=0,j_0=1)}}$ the estimate for case 1 which is depicted by a red line, $\hat{x}_{i,e_{(j_s=1,j_0=0)}}$ the estimate for case 2 which is depicted by a green line and $\hat{x}_{i,e_{(j_s=1,j_0=2)}}$ for case 3 which is depicted by a purple line.
To further show the performance of the proposed observer, additional Gaussian disturbances are applied to the output measurements at time $t = 2s$ with zero mean and standard deviation of 0.1. The simulation results are given in Figure 4.5 with a zoom of the plots between $t = 3s$ and $t = 3.5s$.

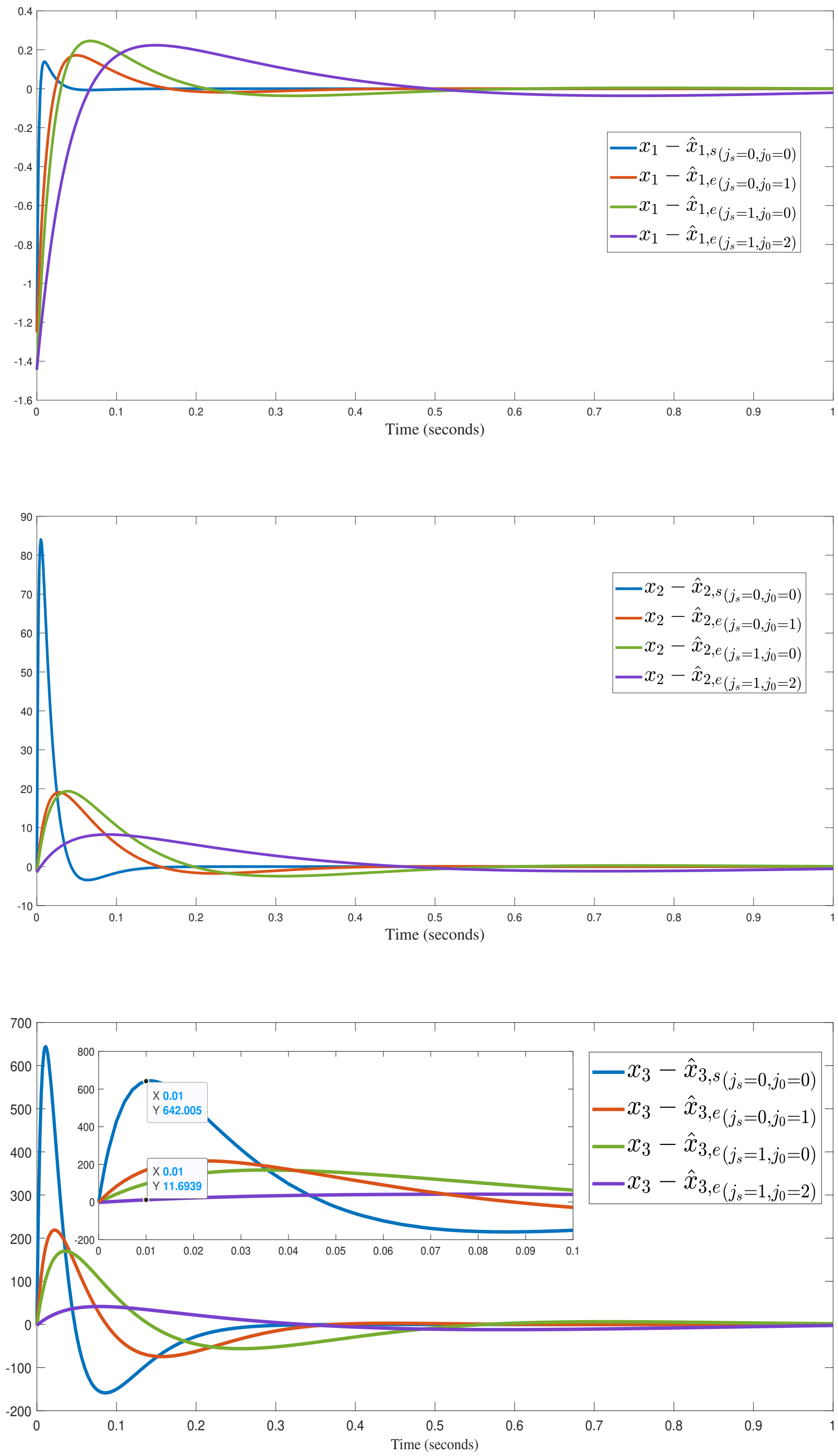



Figure 4.4: Evolution of the observers' error dynamics.

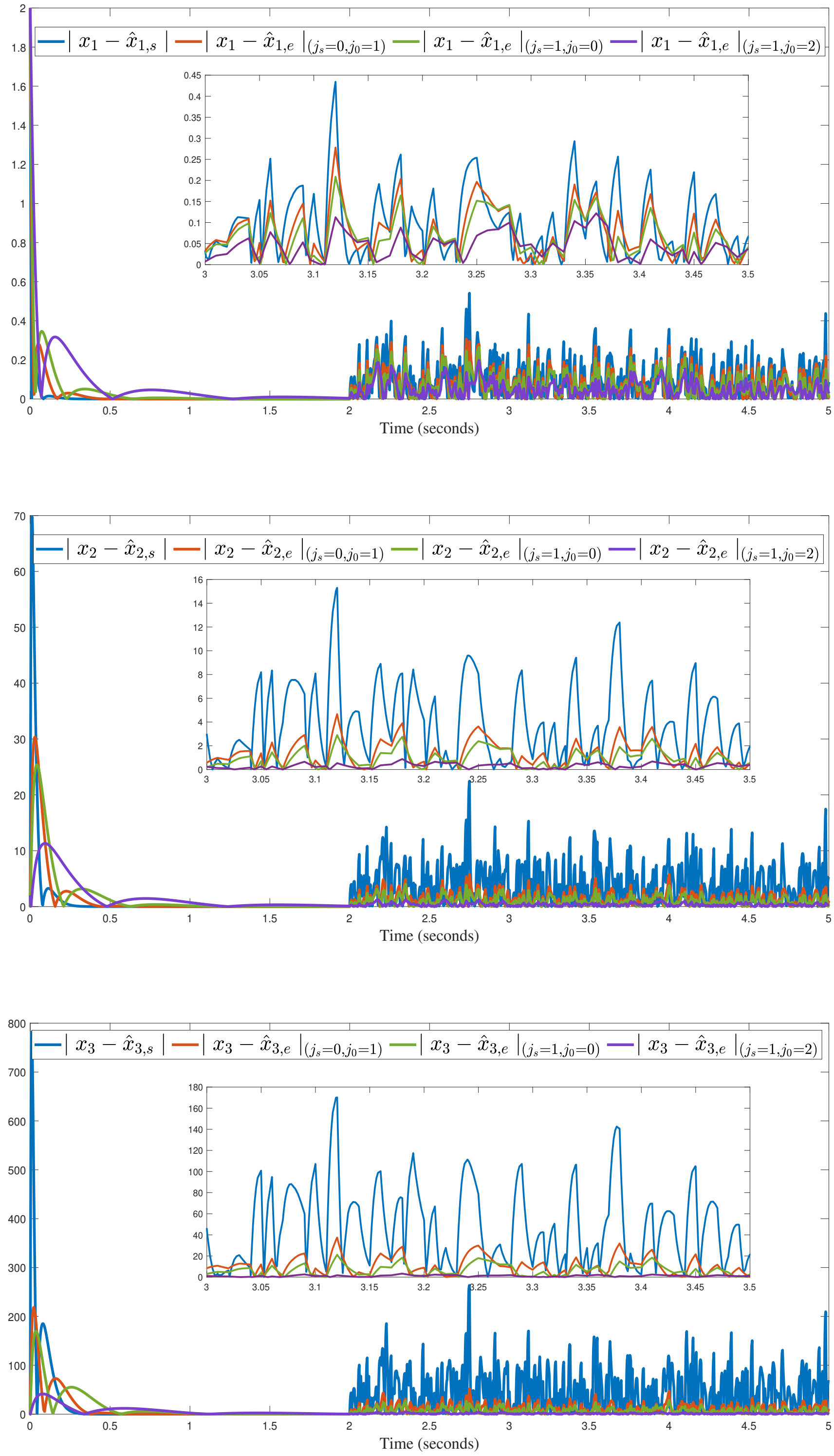



Figure 4.5: Absolute values of estimation errors with additional measurement noise.

## 4.3 SI epidemic model

The spatial spread of infectious diseases, following their introduction at distinct locations, has always been a major concern for human populations due to their high death mortality in both developed and developing countries. About 31.1 million to 43.9 million were living with HIV in 2017 [211], while seasonal influenza epidemics result in about 250,000 to 500,000 deaths worldwide every year, according to the World Health Organization [212]. These examples highlight the significant role of epidemiological models in analyzing the origins, dynamics, and spread of such epidemics. These models provide notation, concepts, intuition, and disease-related factors such as the infectious agent, mode of transmission, latent and infectious periods, susceptibility, and resistance which can be used to capture features that are most influential in the spread of diseases. Such models heavily rely on a good estimation of the epidemiological parameters for simulating the outbreak trajectory and comparing the necessary vaccination levels for herd immunity to make predictions of the disease progression and suggest possible control strategies. Epidemiological constraints, such as delays in symptom appearance (due to incubation period) and positive test confirmation (due to limited testing and detection resources), may limit the real-time use of epidemiological models [213,214]. In order to overcome such constraints, mathematical modeling of infectious diseases was employed in epidemiology, as recognized by WHO [215] and proven to be effective [216,217]. Compartmental modeling as a class of mathematical modeling is often employed in studying the probable outbreak growth [218]. These models consist of two parts: compartments and rules. The compartments divide the population into different possible states with respect to the disease. The rules specify the proportion of individuals moving from one class to another.

Compared to the agent-based model, the compartmental model is simple and correlated to observation [219]. It involves a dynamic process based on how the population is divided into different compartments to describe the transmission state [220]. The SIR (Susceptible, Infectious, and Recovered) model introduced in [221] divides the population into the susceptible, infectious, and recovered compartments to describe the state of disease spread, where people who are susceptible to infection will possibly be infected, and the infected people will be recovered with a certain rate. Different model extensions were developed in the recent past, mostly by including additional compartments, such as the SEIR (Susceptible, Exposed, Infectious and Recovered) model which introduces the exposed compartment into the SIR model to describe the intermediate state between the susceptible and infected people.

Besides the various modeling assumptions underlying the derivation of these models, their practical use relies on two other assumptions: the model parameters are known and an appropriate initial condition, i.e., the current state of the population is known [222]. The problem of estimating the model parameters is certainly important but will not be addressed in this work, one can see for example [223], [224] and [225]. Once the model parameters are known, we can identify the current population state. In most classical models the total population is conserved and this yields an estimate of one of the compartments in terms of estimates of the remaining ones.

Although the literature on the behaviors of epidemic models endowed with a treatment function is vast, there are fewer works on these models from the point of view of observer or estimation theory which allows to predict and control of the propagation of diseases and virus mutation [226]. Earlier work using observers in an epidemiological context dates back to the works in [227] and since 2012 there was a growing interest in the literature such as the estimation of sequestered infected erythrocytes in Plasmodium falciparum malaria patients in [228], parameters and states estimation for an SI-SI Dengue epidemic model [229], interval observer for uncertain SIR and

SIIR-SI models [226], a state observer for a continuous and discrete time SEIR model whose values are then used to implement a vaccination strategy [230].

### 4.3.1 Mathematical Modeling

The SI model splits the population into two groups, the susceptible individuals who may contract the disease and the infected individuals who may spread the disease to the susceptible. Once a susceptible becomes infected, he or she moves into the infected group, increasing the size of the infected class and decreasing the size of the susceptible class as described in the flow diagram depicted in Figure 4.6. In SI modeling some assumptions are made about the population and disease, first, each person in the susceptible population is assumed to be equally likely to transmit the disease through contact with an infected individual, and once a person is infected, they cannot recover; they remain in the infected class forever [231]. Second, the length of the disease outbreak is short compared with the average personâ€™s lifespan, so death is not a factor [232]. Therefore, this model can be applied to diseases for which individuals never recover and for which disease spread is relatively quick, such as herpes (HSV-1 or HSV-2) caused by the virus Herpesviridae. During an epidemic, the population is divided into two compartments: the healthy individuals likely to catch the disease and the infected ones, denoted by $S$ and $I$, respectively, and the total population is represented by $N$.

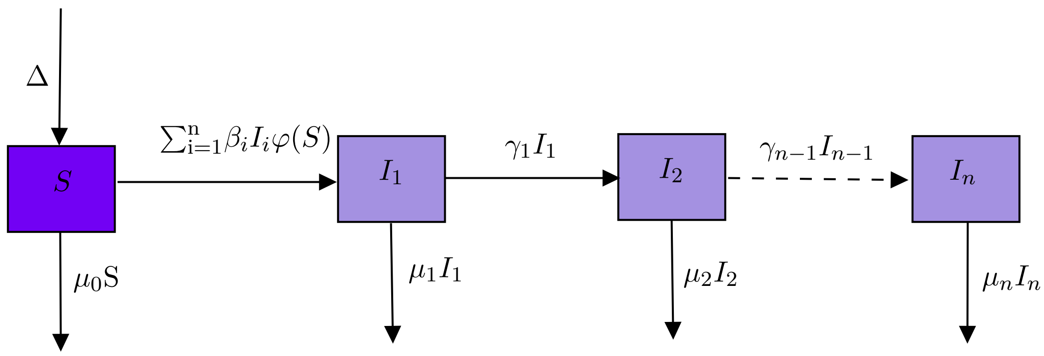



Figure 4.6: The transfer diagram for SI model.

Let us consider a class of stage-structured SI model with $n$ infectious stages which is described by the following dynamics:

$$\dot{S} = \Lambda - \sum\nolimits_{i=1}^{n} \beta_i I_i \varphi(S) - \mu_0 S \tag{4.7}$$

$$\dot{I}_1 = \sum\nolimits_{i=1}^{n} \beta_i I_i \varphi(S) - (\mu_1 + \gamma_1) I_1 \tag{4.8}$$

$$\dot{I}_2 = \gamma_1 I_1 - (\mu_2 + \gamma_2) I_2 \tag{4.9}$$

$$\vdots \tag{4.10}$$

$$\dot{I}_{n-1} = \gamma_{n-2} I_{n-2} - (\mu_{n-1} + \gamma_{n-1}) I_{n-1} \tag{4.11}$$

$$\dot{I}_n = \gamma_{n-1} I_{n-1} - \mu_n I_n \tag{4.12}$$

where $\Delta$ is the recruitment and $\beta_i$ is the per capita contact in the compartment $I_i$. The function $\varphi$ is assumed to be continuous, positive, and increasing that models the exposure of susceptible individuals to contacts with infectious ones, for instance, one can use $\varphi(S) = S^p$ or $\varphi(S) = \frac{S}{1+aS}$

(with $a > 0$) to take into account saturation effects. The parameter $\mu_0$ is the natural death rate of the susceptible individuals and $\mu_i$ is the death rate of the infected individuals in stage $i$, in general, $\mu_i = \mu_0 + d_i$ with $d_i$ being the additional disease-induced mortality rate and $\gamma_i$ is the transition rate from stage $i$ to $i+1$.
Let $x(t) = (S(t), I_1(t), \dots, I_n(t))^T$ be the state vector of the system (4.7) and we suppose that we can only measure the level of infection in the last stage, i.e., the measurable output of the system is $y(t) = I_n(t)$. The aim of this work is: knowing $I_n(t) \forall t \geq 0$, we need to derive estimates $\hat{S}(t)$ and $\hat{I}_i(t)$ satisfying $\lim_{t\to\infty} \hat{S}(t) - S(t) = 0$ and $\lim_{t\to\infty} \hat{I}(t) - I(t) = 0$ using the high-gain observer techniques introduced in Chapter 2.

### 4.3.2 Two stage structure SI model

In this section, we consider the special case where compartment $I$ is made up of two compartments, namely: the infected in the first stage of the disease and the infected in the terminal phase of the infection, denoted $I_1$ and $I_2$, respectively. The corresponding SI model is given as follows,

$$\Sigma_2 : \begin{cases} \dot{S} = \Delta - (\beta_1 I_1 + \beta_2 I_2)S - \mu_0 S, \\ \dot{I}_1 = (\beta_1 I_1 + \beta_2 I_2)S - (\mu_1 + \gamma)I_1, \\ \dot{I}_2 = \gamma I_1 - \mu_2 I_2 \end{cases} \tag{4.13}$$

where

- $S, I_1, I_2$ represent the compartments of susceptible, first-stage infected, and second-stage infected, respectively.

- $\Delta, \mu, \alpha, 1/\gamma$ are recruitment, mortality rate, recovery rate, and time taken for an early-stage infected to become in the final phase of infection, respectively.

Let $x(t) = (S(t), I_1(t), I_2(t))$ be the state vector of the system (4.13) and we suppose that we can only measure the level of infection in the last stage, i.e., the measurable output of the system is $y = I_2$. The aim of this work is to derive estimates $\hat{S}(t)$ and $\hat{I}_i(t)$ using high-gain observer techniques.
The model given by (4.13) has the following form:

$$\begin{cases} \dot{x} = f(x) \\ y = h(x) \end{cases} \tag{4.14}$$

in which $x \in \mathbb{R}^3$, hence there is a "physical subset" $\Omega \subset \mathbb{R}^3$ where the system lies. Let us construct the $j^{th}$ time derivative of the output. This can be expressed using Lie differentiation of the function $h$ by the vector field $f$, $L_f^j(h)(x(t))$ given as

$$\begin{aligned} L_f^0(h)(x) &= h(x), \\ L_f^j(h)(x) &= \frac{\partial}{\partial x}(L_f^{j-1}(h)(x))f(x). \end{aligned} \tag{4.15}$$

When (4.13) is observable, the map $\Psi : x \to \Psi(x)$ is a diffeomorphism and,

$$z = \Psi(x) = \begin{bmatrix} I_2 \\ \gamma I_1 - \mu_2 I_2 \\ \gamma(\beta_1 I_1 + \beta_2 I_2)S - \gamma(\mu_1 + \mu_2 + \gamma)I_1 + \mu_2^2 I_2 \end{bmatrix}, \tag{4.16}$$

its Jacobian is given as

$$\frac{\partial \Psi}{\partial x} = \begin{bmatrix} 0 & 0 & 1 \\ 0 & \gamma & -\mu_2 \\ \gamma\left(\beta_1 I_1 + \beta_2 I_2\right) & \gamma(\beta_1 S - m) & \gamma\beta_2 S + \mu_2^2 \end{bmatrix}. \tag{4.17}$$

The determinant of the Jacobian is equal to $det\left(\frac{\partial \Psi}{\partial x}\right) = \gamma^2(\beta_1 I_1 + \beta_2 I_2)$, which is positive in the positively invariant open set $\Omega_0 = \left\{S > 0, I_1 > 0, I_2 > 0, S + I_1 + I_2 < \frac{\Delta}{\mu_0}\right\}$, see [233].
With the change of variable $z = \Psi(S, I_1, I_2)$, the expression of $\Psi^{-1}$ is

$$\Psi^{-1} : z \to \begin{cases} \dfrac{\mu_2\left(\mu_1 + \gamma\right) z_1 + \left(\mu_1 + \mu_2 + \gamma\right) z_2 + z_3}{\left(\mu_2\beta_1 + \gamma\beta_2\right) z_1 + \beta_1 z_2}, \\ \frac{\mu_2}{\gamma} z_1 + \frac{1}{\gamma} z_2 \\ z_1 \end{cases} \tag{4.18}$$

Then, the system given by (4.13) is rewritten in the following triangular form:

$$\begin{cases} \dot{z} = F'(z) = \begin{bmatrix} \dot{z}_1 \\ \dot{z}_2 \\ \dot{z}_3 \end{bmatrix} = \begin{bmatrix} z_2 \\ z_3 \\ \varphi(z) \end{bmatrix}, \\ y = Cz = \begin{bmatrix} 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} z_1 \\ z_2 \\ z_3 \end{bmatrix} \end{cases} \tag{4.19}$$

where $\varphi$ can be extended from $\Omega$ to the entire $\mathbb{R}^3$ by a $\mathcal{C}^\infty$ function globally Lipschitz on $\mathbb{R}^3$ and according to the expression of $\Psi^{-1}$ we deduce $\varphi(z)$ as given in (4.20).

$$\begin{aligned} \varphi(z) = &\Big(\left(\beta_1\mu_2 + \beta_2\gamma\right) z_2 + \beta_1 z_3\Big) \frac{\mu_2\left(\mu_1 + \gamma\right) z_1 + \left(\mu_1 + \mu_2 + \gamma\right) z_2 + z_3}{\left(\beta_1\mu_2 + \beta_2\gamma\right) z_1 + \beta_1 z_2} \\ &+ \Big(\left(\beta_1\mu_2 + \beta_2\gamma\right) z_1 + \beta_1 z_2\Big) \Big(\Delta - \frac{\left(\mu_1 + \gamma\right)\mu_2}{\gamma} z_1 - \frac{\mu_1 + \mu_2 + \gamma}{\gamma} z_2 - \frac{z_3}{\gamma} \\ &-\mu_0 \frac{\mu_2\left(\mu_1 + \gamma\right) z_1 + \left(\mu_1 + \mu_2 + \gamma\right) z_2 + z_3}{\left(\beta_1\mu_2 + \beta_2\gamma\right) z_1 + \beta_1 z_2}\Big) - \left(\mu_1 + \gamma\right) z_2 - \left(\mu_1 + \mu_2 + \gamma\right) z_3 \end{aligned} \tag{4.20}$$

### 4.3.3 Simulation results

This latter form allows us to make use of the high-gain observer techniques presented in the previous section to estimate the evolution of susceptible individuals $S(t)$ and infected individuals $I(t)$. The parameters used in these simulations are given as follows: $\Delta = 40$, $\beta_1 = 0.01$, $\beta_2 = 0.15$, $\mu_0 = \mu_1 = 0.01$, $\mu_2 = 0.025$ and $\gamma = 0.02$.
In order to test the efficiency of the proposed observer given in Chapter 2, we performed a batch of numerical simulations to compare with the standard high-gain observer in a noise-free case and in the presence of high-frequency sensor noise, numerically taken as Gaussian distributed noise with zero mean and standard deviation of 0.1. First, we design a standard high-gain observer for the system (4.19) following Theorem 4 to obtain the values of the gain $K$ and the

tuning parameter $\theta$ as given in Table 4.3. Next, following the design methodology described in Chapter 2, we obtain the values of the corresponding gain $K_\Gamma$ and the new observer parameter $\theta_\Gamma$ for different values of the indexes $j_s$ and $j_0$ as summarized in Table 4.3.

Table 4.3: Comparison between the high-gain observers for the SI epidemic model.

| **Standard high-gain observer** | | | | |
|---|---|---|---|---|
| $n_{\text{LMI}}$ | $\theta$ | $K$ | | |
| 1 | 43.46 | 11.67 | 20.97 | 10.53 |

| **Enhanced high-gain observer** | | | | | | |
|---|---|---|---|---|---|---|
| $j_s$ | $j_0$ | $n_{\text{LMI}}$ | $\theta_\Gamma^{\frac{1}{(1+j_s)(1+j_0)}}$ | $K_\Gamma$ | | |
| 0 | 0 | 1 | 43.46 | 11.67 | 20.97 | 10.53 |
| | 1 | 2 | 8.09 | 12.49 | 26.60 | 23.24 |
| | 2 | 4 | 6.75 | 17.53 | 61.13 | 98.10 |

| **Enhanced high-gain observer** | | | | | | | |
|---|---|---|---|---|---|---|---|
| $j_s$ | $j_0$ | $n_{\text{LMI}}$ | $\theta_\Gamma^{\frac{1}{(1+j_s)(1+j_0)}}$ | $K_\Gamma$ | | | |
| 1 | 0 | 1 | 10.11 | 8.87 | 15.79 | 29.16 | 18.27 |
| | 1 | 2 | 3.31 | 10.03 | 27.34 | 33.16 | 21.27 |
| | 2 | 4 | 2.46 | 12.13 | 35.63 | 53.13 | 43.08 |

In Figure 4.7, the estimation error dynamics $(x_i - \hat{x}_i)$ are plotted for $i = 1 \ldots 3$, where $\hat{x}_{i,s}$ denotes the state estimate using the standard high-gain observer which is represented in the blue line, $\hat{x}_{i,e_{(j_s=0,j_0=1)}}$ the state estimate using the HG/LMI technique which represented with a red line, $\hat{x}_{i,e_{(j_s=1,j_0=0)}}$ using the system state augmentation approach which is represented with a green line, and $\hat{x}_{i,e_{(j_s=1,j_0=2)}}$ using the combined HG/LMI and the state augmentation approach represented by the purple line.

For robustness comparison, we have analyzed the behavior of the standard high-gain observer and the new proposed observer for three different cases corresponding to ($j_s = 0$, $j_0 = 1$), ($j_s = 1$, $j_0 = 0$) and ($j_s = 1$, $j_0 = 2$) under noisy outputs measurements. The simulations are done using an additive noise measurement at $t = 2s$, which is a Gaussian-distributed random signal with zero mean and standard deviation of 0.1, the results are shown in Figure 4.8.

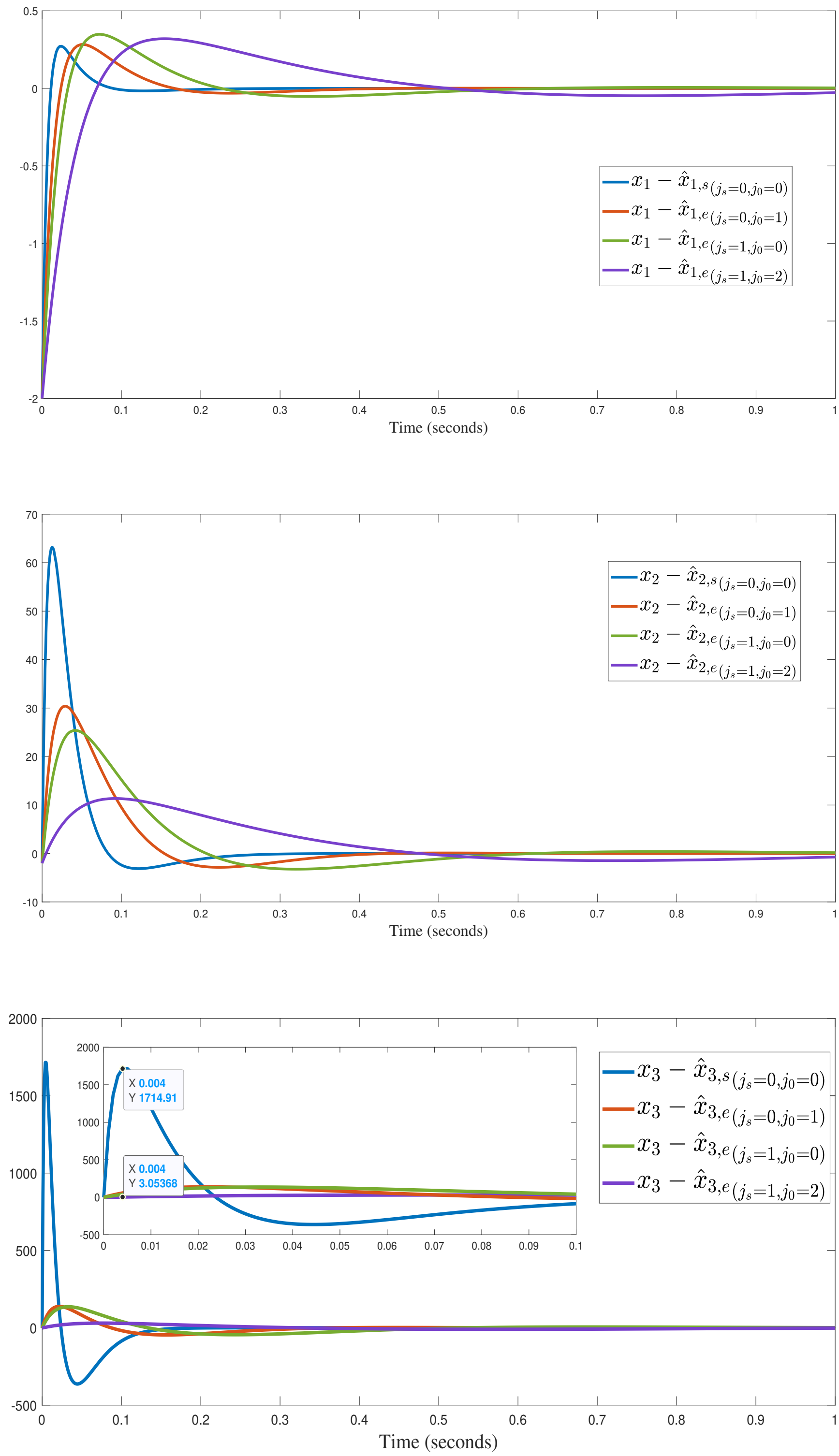



Figure 4.7: Evolution of the observers' error dynamics.

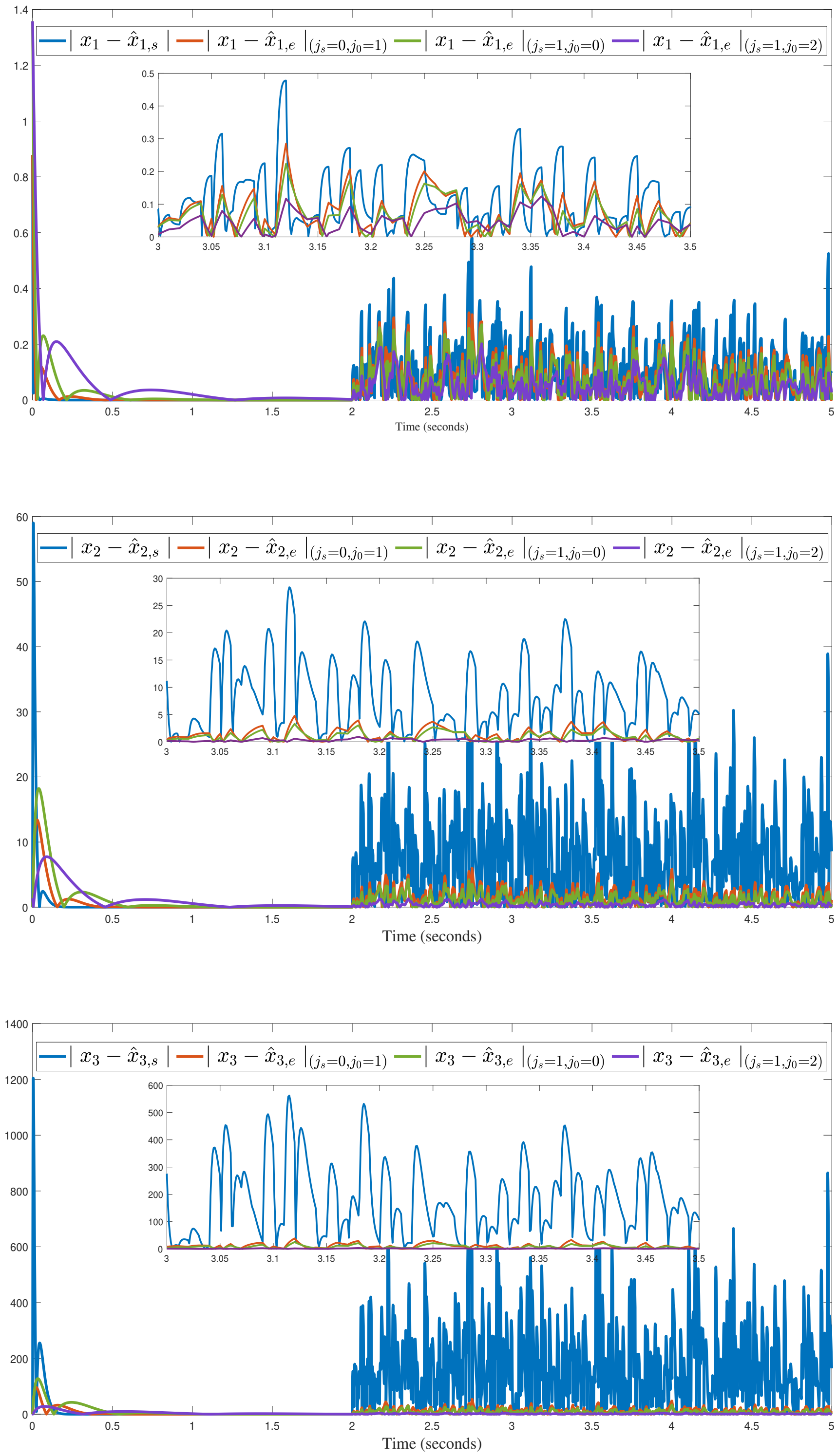



Figure 4.8: Absolute values of estimation errors with additional measurement noise.

## 4.4 Amnioserosa Cell Dynamics During Drosophila Dorsal Closure

The dynamic behavior of epithelial cell sheets plays a central role in morphogenesis, wound healing, tumor growth, tissue engineering, and other biological processes. These processes occur as a result of cell adhesion, migration, division, differentiation, and death, and involve multiple processes acting at the cellular and molecular level [234].
The process of Dorsal Closure (DC) in Drosophila melanogaster represents a good model system for the study of epithelial sheet morphogenesis which has long been of interest due to its experimental accessibility [235–238]. It is a complex morphogenetic process that results in the closing of a gap that exists halfway through embryogenesis in the dorsal epidermis [239]. During the closure, lateral epidermal cells elongate along the dorsoventral axis and subsequently spread dorsally to cover the embryonic dorsal surface [240], [241]. On the other hand, cell shape oscillations occur in the Amnioserosa cells on the dorsal surface as they contract and eventually disappear inside the embryo so that the two opposing epithelial fuse at the dorsal midline to establish epidermal continuity; thus, the increase in epidermal surface area is accommodated by a reduction in the Amnioserosal surface area [236, 242–244].
Aside from its biological significance, DC of Drosophila provides a fascinating example of tissue morphogenesis that couples cell and tissue mechanics with intra- and intercellular structural remodeling which made it the target of a number of mathematical models that seek to explain and quantify the processes that underlie closure's kinematics [245]. To this end, a variety of different cell-based modeling approaches have been developed, which treat cells as individual objects allowing the study of cellular processes such as motility, adhesion, mitosis, and apoptosis and their effect on the dynamic behavior of epithelia (for a review, see [245]). The application of these mathematical models is useful to provide a handy tool for estimating both the mechanical and biochemical properties of epithelial morphogenesis in a non-invasive manner. They are also necessary to determine the values of some physical parameters that are not measurable during experiments such as moduli of diffusion or elasticity.
Considering that mechanical characteristics are very difficult to probe in vivo, most of the biomechanical studies are performed in vitro, on single molecules [246], single cells [247], or minimal systems such as in-vitro grown molecular networks [248]. Laser perturbations can also provide valuable information [249, 250], although they can be very damaging to the studied organism. In view of this, numerical studies and nonlinear estimation methods can be useful to provide handy tools for estimating the mechanical properties of biological systems in a non-invasive manner. In the past few years, several works focused on forces and deformations in live tissues [251, 252]; yet, none of them allows one to study the parameter variations over time, and neither do they permit to assess the mechanical characteristics of a tissue.
In line with some previous work [193], the mechanical model of live Aamnioserosa cells during Drosophila melanogaster's dorsal closure will be investigated in this work by considering mass-spring vertex model. Using such a model we will estimate the cellâ€™s vertex positions and the parameter $\kappa$ which stands from the spring constants between vertex-vertex and vertex-barycenter connection in order to track the cell's movements.

### 4.4.1 Amnioserosa cells oscillation model

The role of Amnioserosa cells (AS) in Dorsal Closure, and the mechanics of these cells, have been the objects of several studies, in particular how their apical area oscillations influence closure. Yet, no consensus arose from these studies to explain the role of AS cells. In this section, we

investigate the mechanics behind the AS cell oscillations and build a dynamic model representing these movements based on a mass-spring lattice model.
Amnioserosa cells are known to show a periodic oscillation of their apical area [253]. These oscillations are due to the periodic variation of the concentration of non-muscular myosin II at the center of the apical part of the cell. Myosin II is a molecular motor, binding to actin fibers and generating a contractile force. Hence, a natural way to represent Amnioserosa cells' movement is to consider them a set of autonomous oscillators interacting with each other. In this work, we consider the vertices (junctions between three cells) to be focal points for forces which gives rise to a lattice model whose nodes positions are the vertices. The Amnioserosa cells are also capable to contract and expand radially in a periodic manner, hence, the cells' barycenters are also included in the lattice as shown in figure 4.9.

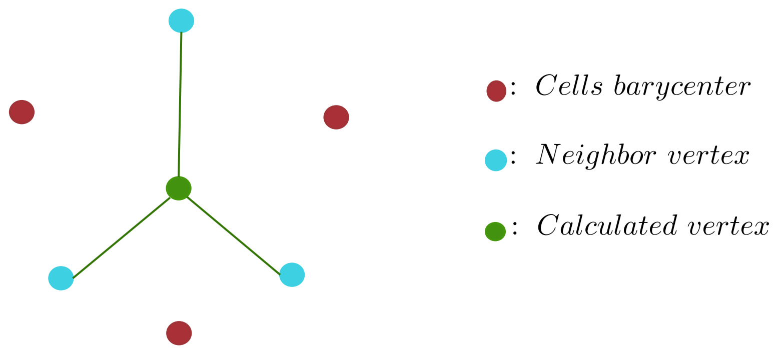



Figure 4.9: Schematized cellular junctions.

In the case of the pure elastic model, the position of the central (green) vertex satisfies the following equation (4.21):

$$\ddot{q}(t) = -\frac{K_{\text{vertex}}}{m} \sum_{j=1}^{n_{\text{vertex}}} (q_i - q_j - l_{0,ij}) \tag{4.21}$$

where, $l_{0,ij}$ represents the resting length between vertex $i$ and its neighbors $j = 1, 2, 3$, $q_i$ and $q_j$ are the positions of the calculated vertex and its neighbors, respectively.

### 4.4.2 Mathematical modeling

To represent the cellular dynamics, we rely on a periodic hexagonal lattice model as depicted in Figure 4.10. A lattice model is in general composed of $n$ cells with $n_v$ vertices, which are junctions connecting cells. Such vertices are regarded as the focal points of forces in the tissue and are represented by point masses of cartesian $2D$ coordinates (with respect to a fixed frame) denoted by $q_i(t) \in \mathbb{R}^2$ at time $t \geq 0$ and $i = 1, 2, \ldots n_v$ with $q_{i,j}$ denoting the $j^{th}$ component of the coordinate of the mass point $i$. We assume that the mass of each of such points is $m$ and the velocity of the mass point $i$ at time $t$ is denoted by $\dot{q}_i(t) \in \mathbb{R}^2$. The mechanical parameters are divided into two distinct sets, which account for the radial and azimuthal contributions of the cells. The mass points are assumed to be the same for every vertex. The model requires the introduction of the cell barycenters, each of which is without mass and with coordinate $r_j(t) \in \mathbb{R}^2$ known upon measurements for all $j = 1, 2, \ldots, n$. To complete the model, we need in general to introduce the set of neighbor vertices and barycenters of a given vertex $i$, which will be denoted by $V_i$ and $B_i$, respectively, then, using the Lagrangian approach with a specific energy function, we will analyze the linear elastic model (LEM).

In the LEM, the total energy associated with the point mass $i$ is:

$$L_{LEM_i} = \frac{1}{2}m\dot{q}_i(t)^2 - \frac{K_v}{2}\sum_{j\in V(i)}(q_i(t) - q_j(t) - l_{1,i,j})^2 - \frac{K_b}{2}\sum_{j\in B(i)}(q_i(t) - r_j(t) - l_{2,i,j})^2 \quad (4.22)$$

where $K_v$ and $K_b$ are the azimuthal (vertex-vertex connectivity) and radial (vertex-barycenter connectivity) elastic moduli respectively, $l_{1,i,j}$ and $l_{2,i,j}$ the resting lengths of the springs between vertex $i$ and another vertex or barycenter $j$, respectively.

Based on (4.22), we obtain the Euler-Lagrange equations:

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_{i,j}}\right) - \frac{\partial L}{\partial q_{i,j}} = 0, i = 1, 2, \ldots, n_v, j = 1, 2 \quad (4.23)$$

where $L = \sum_{i=1}^{n_v} L_{LEM_i}$, is the Lagrangian.

The resulting $2n_v$ equations can be written as follows:

$$\ddot{q}_i(t) = \frac{K_v}{m}\sum_{j\in V(i)}(q_i(t) - q_j(t) - l_{1,i,j}) + \frac{K_b}{m}\sum_{j\in B(i)}(q_i(t) - r_j(t) - l_{2,i,j}) \quad (4.24)$$

The cell barycenters $r_j \in \Re^2$ are without mass and are known upon measurements for all $j = 1, 2, \ldots n$.

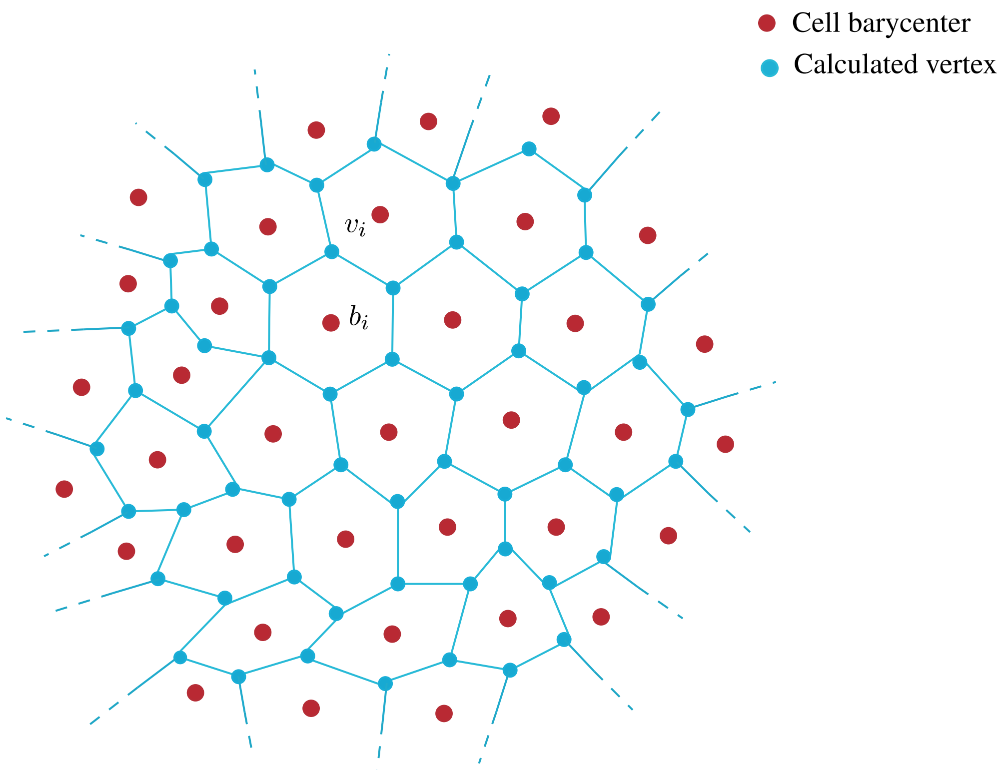



Figure 4.10: Hexagonal periodic system of cells: blue and red nodes indicate cell vertices and barycenters, respectively. Blue lines indicate apical junctions between adjacent cells.

To address the cellular dynamics during dorsal closure, we have selected one cell from the system of cells depicted in Figure 4.10 and applied the LEM to study the cellular movement and evolution over time. The cell is represented by six nodes (vertices) on the periphery and one central node (barycenter), each vertex is connected to its neighboring vertex or barycenter with an elastic latice which is represented by a spring of stiffness $K$ as shown in Figure 4.11.

$$
\begin{array}{llll}
V(v_1) = \{v_2, v_7, v_6\}, & V(v_7) = \{v_1\}; & B(v_1) = \{b_1, b_6, b_7\}, & B(v_7) = \{b_6, b_7\}; \\
V(v_2) = \{v_1, v_3, v_8\}, & V(v_8) = \{v_2\}; & B(v_2) = \{b_1, b_2, b_7\}, & B(v_8) = \{b_2, b_7\}; \\
V(v_3) = \{v_2, v_4, v_9\}, & V(v_9) = \{v_3\}; & B(v_3) = \{b_1, b_2, b_3\}, & B(v_9) = \{b_2, b_3\}; \\
V(v_4) = \{v_3, v_5, v_{10}\}, & V(v_{10}) = \{v_4\}; & B(v_4) = \{b_1, b_3, b_4\}, & B(v_{10}) = \{b_3, b_4\}; \\
V(v_5) = \{v_4, v_6, v_{11}\}, & V(v_{11}) = \{v_5\}; & B(v_5) = \{b_1, b_4, b_5\}, & B(v_{11}) = \{b_4, b_5\}; \\
V(v_6) = \{v_1, v_5, v_{12}\}, & V(v_{12}) = \{v_6\}; & B(v_6) = \{b_1, b_5, b_6\}, & B(v_{12}) = \{b_5, b_6\}.
\end{array}
$$

$v_i$ and $b_i$ represent the vertices and barycenters, respectively.

Using equation (4.24), the dynamic equation of each vertex is written to get the following set of equations:

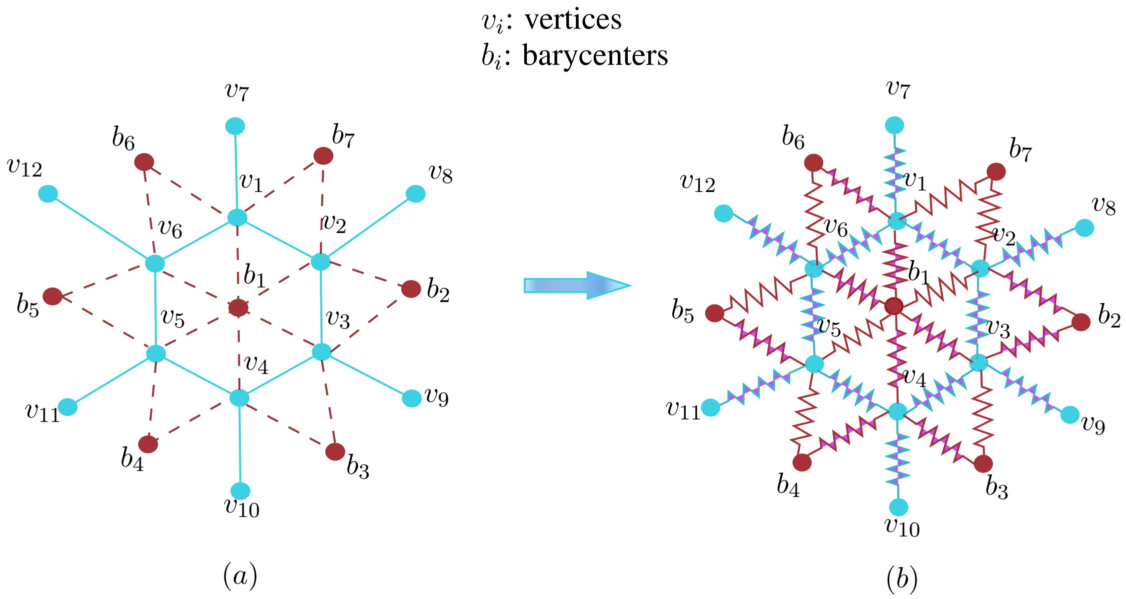



Figure 4.11: (a) Lattice model representation of one single cell with its adjacent vertices and barycenters. (b) Spring model representation: blue springs connect adjacent vertices (cortical visco-elasticity) and red springs connect vertices and cell barycenters (medioapical array visco-elasticity).

$$\ddot{q}_1 = \frac{3(K_v + K_b)}{m} q_1 - \frac{K_v}{m}(q_2 + q_6 + q_7) - \frac{K_b}{m}(b_1 + b_6 + b_7) \tag{4.25a}$$

$$\ddot{q}_2 = \frac{3(K_v + K_b)}{m} q_2 - \frac{K_v}{m}(q_1 + q_3 + q_8) - \frac{K_b}{m}(b_1 + b_2 + b_7) \tag{4.25b}$$

$$\ddot{q}_3 = \frac{3(K_v + K_b)}{m} q_3 - \frac{K_v}{m}(q_2 + q_4 + q_9) - \frac{K_b}{m}(b_1 + b_2 + b_3) \tag{4.25c}$$

$$\ddot{q}_4 = \frac{3(K_v + K_b)}{m} q_4 - \frac{K_v}{m}(q_3 + q_5 + q_{10}) - \frac{K_b}{m}(b_1 + b_3 + b_4) \tag{4.25d}$$

$$\ddot{q}_5 = \frac{3(K_v + K_b)}{m} q_5 - \frac{K_v}{m}(q_4 + q_6 + q_{11}) - \frac{K_b}{m}(b_1 + b_4 + b_5) \tag{4.25e}$$

$$\ddot{q}_6 = \frac{3(K_v + K_b)}{m} q_6 - \frac{K_v}{m}(q_1 + q_5 + q_{12}) - \frac{K_b}{m}(b_1 + b_5 + b_6) \tag{4.25f}$$

Note that for the sake of brevity, the terms $l_{1,i,j}$ and $l_{2,i,j}$ are deleted from (4.25) but are taken into consideration in the simulation part.
Then, from (4.25) we get the following first order system
the dynamic equation of each vertex is written to get the following first-order system:

$$\begin{cases} \dot{q}_i(k) = A(\kappa)\, q_i(k) + B(\kappa)\, u(k) \\ y(k) = C\, q(k) \end{cases} \tag{4.26}$$

where: $\kappa = \begin{bmatrix} \frac{K_v}{m} & \frac{K_b}{m} \end{bmatrix}^\top$, and the matrices $A$, $B$ and $C$ are given as follows:

$$A = \begin{bmatrix} A_{(1,1)} & A_{(1,2)} & \cdots & \cdots & A_{(1,11)} & A_{(1,12)} \\ A_{(2,1)} & A_{(2,2)} & \cdots & \cdots & A_{(2,11)} & A_{(2,12)} \\ \vdots & \vdots & \ddots & \ddots & \vdots & \vdots \\ \vdots & \vdots & \ddots & \ddots & \vdots & \vdots \\ A_{(11,1)} & A_{(11,2)} & \cdots & \cdots & A_{(11,11)} & A_{(11,12)} \\ A_{(12,1)} & A_{(12,2)} & \cdots & \cdots & A_{(12,11)} & A_{(12,12)} \end{bmatrix} \tag{4.27}$$

where $A_{(i,j)}$ are block matrices with $(i = 1, 2, \ldots, 12)$ and $(j = 1, 2, \ldots, 12)$, given as follows:

$$A_{ij} = \begin{cases} \begin{bmatrix} 0 & 1 \\ K & 0 \end{bmatrix}; & \text{for } \begin{cases} i = j \\ K = \frac{3(K_v+K_b)}{m}; & i = 1, \ldots, 6 \\ K = \frac{(K_v+2K_b)}{m}; & i = 7, \ldots, 12 \end{cases} \\ \begin{bmatrix} 0 & 0 \\ \frac{K_v}{m} & 0 \end{bmatrix}; & \text{for } \begin{cases} |i-j| = 6 \end{cases} \\ \begin{bmatrix} 0 & 0 \\ \frac{K_v}{m} & 0 \end{bmatrix}; & \text{for } \begin{cases} i = 1, \ldots, 6 \\ |i-j| = 1 \\ |i-j| = 5 \end{cases} \\ \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}; & \text{elsewhere} \end{cases}$$

$$B = \begin{bmatrix} b_{(1,1)} & b_{(1,2)} & \cdots & \cdots & b_{(1,6)} & b_{(1,7)} \\ b_{(2,1)} & b_{(2,2)} & \cdots & \cdots & b_{(2,6)} & b_{(2,7)} \\ \vdots & \vdots & \ddots & \ddots & \vdots & \vdots \\ \vdots & \vdots & \ddots & \ddots & \vdots & \vdots \\ b_{(23,1)} & b_{(11,2)} & \cdots & \cdots & b_{(23,6)} & b_{(24,7)} \\ b_{(24,1)} & b_{(12,2)} & \cdots & \cdots & b_{(24,6)} & b_{(24,7)} \end{bmatrix} \tag{4.28}$$

with $b_{(i,j)}$ are defined as follows:

$$b_{(i,j)} = \begin{cases} \frac{K_b}{m}; & \text{for } \begin{cases} i = 2; & j = 1, 6, 7 \\ i = 4; & j = 1, 2, 7 \\ i = 6, 8, 12; & j = 1; \quad i - 2j = 0, 2 \\ i = 14; & j = 6, 7 \\ i = 16; & j = 2, 7 \\ i = 20, 22, 24; & i - 2j = 12, 14 \end{cases} \\ 0; & \text{elsewhere} \end{cases}$$

We assume that we can measure the coordinates of the vertices surrounding the cell i.e., $q_7, q_8, q_9, q_{10}, q_{11}, q_{12}$, hence, the matrix $C$ is defined as:

$$C = \begin{bmatrix} c_{(1,1)} & c_{(1,2)} & \cdots & \cdots & c_{(1,23)} & c_{(1,24)} \\ c_{(2,1)} & c_{(2,2)} & \cdots & \cdots & c_{(2,23)} & c_{(2,24)} \\ \vdots & \vdots & \ddots & \ddots & \vdots & \vdots \\ \vdots & \vdots & \ddots & \ddots & \vdots & \vdots \\ c_{(11,1)} & c_{(11,2)} & \cdots & \cdots & c_{(11,23)} & c_{(11,24)} \\ c_{(12,1)} & c_{(12,2)} & \cdots & \cdots & c_{(12,23)} & c_{(12,24)} \end{bmatrix} \tag{4.29}$$

where $c_{(i,j)}$ are defined as:

$$c_{(i,j)} = \begin{cases} 1; & \text{for } \begin{cases} i = 1,3,5,7,9,11 \\ j - i = 12 \end{cases} \\ 0; & \text{elsewhere} \end{cases}$$

After forward Euler discretization of the system described by (4.26) and taking into consideration the system and measurements disturbances ($w_t \in \mathbb{R}^n$) and ($v_t \in \mathbb{R}^m$) respectively, the following discrete-time system is obtained:

$$\begin{cases} q_{t+1} = A_d(\kappa_t)\, q_t + B_d(\kappa_t)\, u_t + w_t \\ y_t = C_d\, q_t + v_t \end{cases} \tag{4.30}$$

with $A_d = I + \Delta\, A$, $B_d = \Delta\, B$ and $\Delta$ is the time discretization step.

### 4.4.3 Simultion results

In this section, we refer to the cell's model described in section 4.4.2, and using the MHE and EKF, we estimate the distance between the vertices $v_1$, $v_2, v_3$, $v_4, v_5$, $v_6$ and the barycenter $b_1$ to track the cell dynamics (contraction and retraction), together with the variable parameters $\frac{K_v}{m}$ and $\frac{K_b}{m}$ over a horizon of 100 steps and a moving window of $N = 10$. The unknown parameters are randomly generated in the interval $[0, 1]$ initially and subjected to an additive, zero-mean Gaussian noise with variance 0.01, lower or upper saturated in case it goes out of the interval $[0, 1]$. The result of a simulation run is depicted in Figures 4.12, 4.16 and 4.13.
Figure 4.12 shows the evolution of the distances between vertices $v_i$ with $i = 1, \dots, 6$ and barycenter $b_1$ with their corresponding estimates using both MHE and EKF. As we can see from the figure, the MHE shows a better accuracy than the EKF which begins to diverge after $t = 50$.
Figure 4.13 depicts the actual and estimated values of the unknown parameters $\frac{K_v}{m}$ and $\frac{K_b}{m}$. The results show that MHE provides better estimates than the EKF as it allows incorporation of the system constraints.
In order to further show the performances of the MHE and the EKF, two criteria are supposed for comparison between these two latters. The first criterion is the Root Mean Squared Error (RMSE), and the second is the computational time representing the simulation execution time for each state estimation method. The results of RMSE for both MHE and EKF are summarized in Table 4.4 and Figure 4.14 where it can be clearly seen that the minimum RMSE belongs to the MHE as it provides the fittest curves with the smallest estimation error.

| Estimator | MSE ($dv_1$) | MSE ($dv_2$) | MSE ($dv_3$) | MSE ($dv_4$) | MSE ($dv_5$) | MSE ($dv_6$) |
|---|---|---|---|---|---|---|
| MHE | 0.1111 | 0.0888 | 0.1242 | 0.1107 | 0.1056 | 0.1896 |
| EKF | 0.4313 | 0.3645 | 0.3909 | 0.3375 | 0.2801 | 0.1311 |

Table 4.4: Medians of RMSEs of a single optimization round for the MHE and EKF.

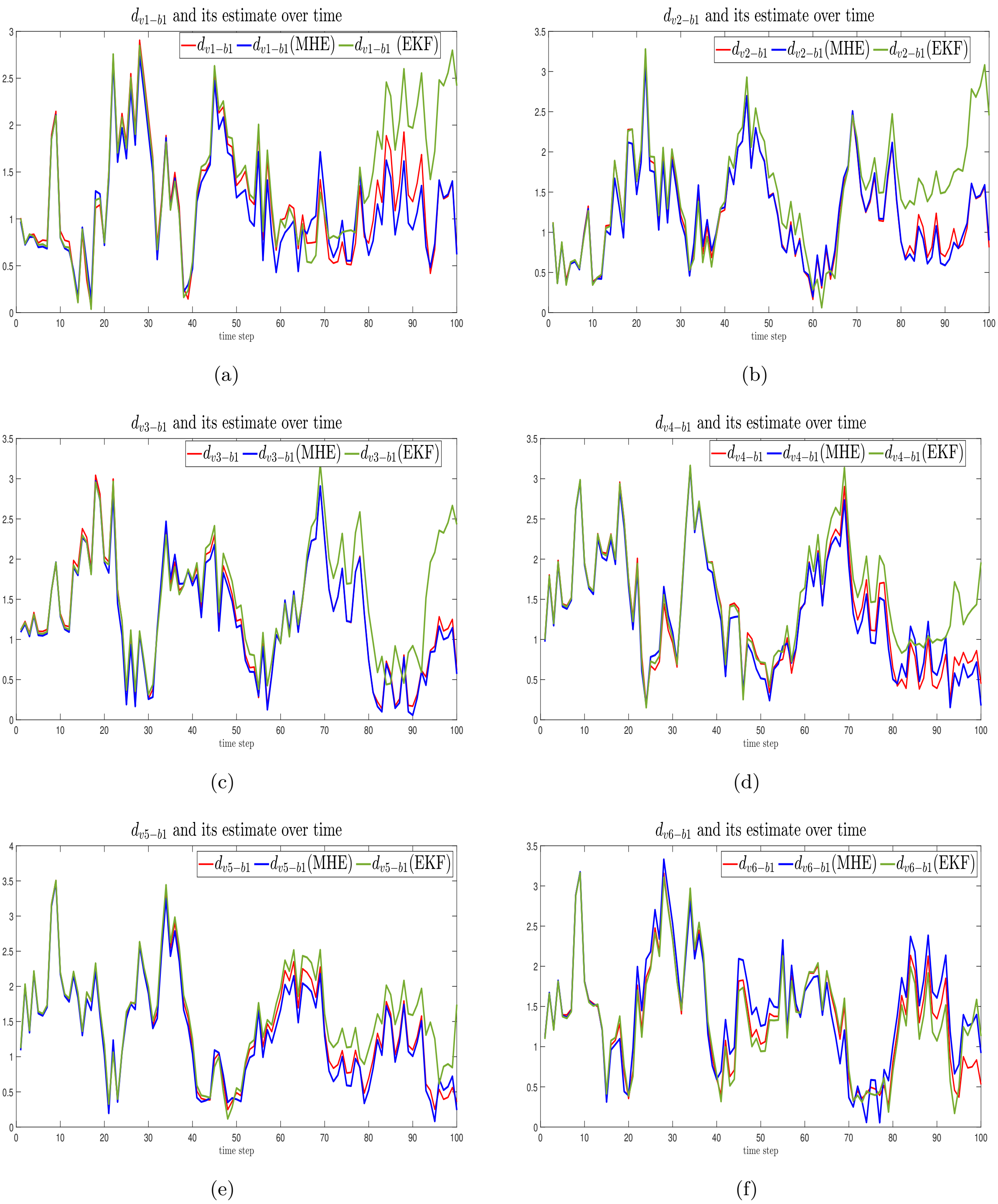
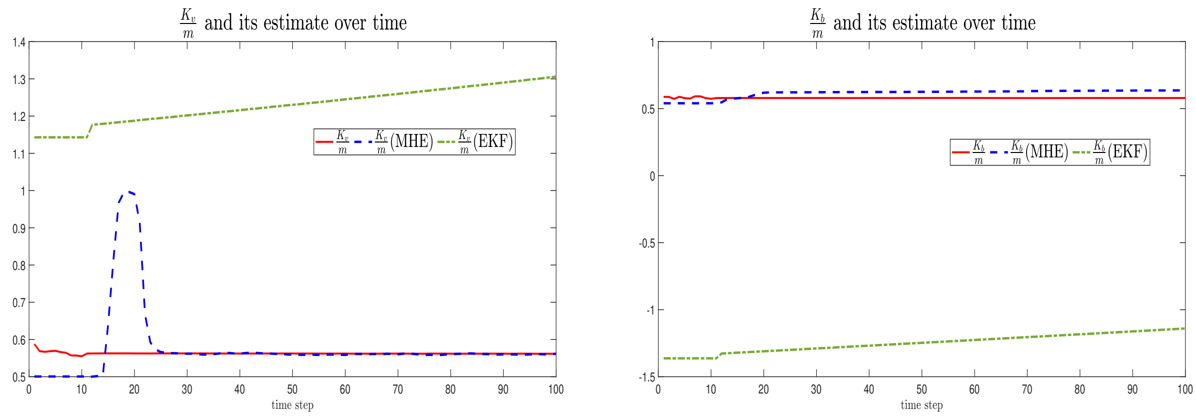
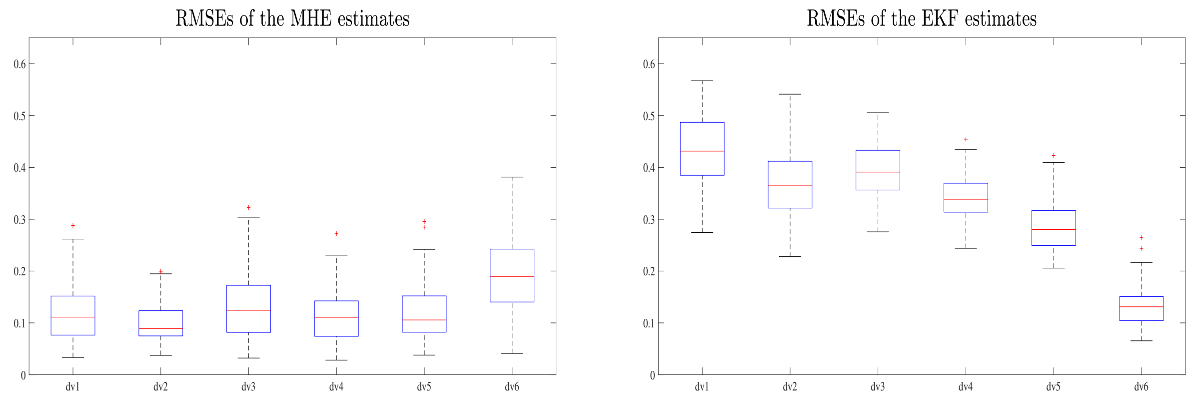



Figure 4.12: Distance between vertex $v_i$ and barycenter $b_1$: (a) $d_{(v_1-b_1)}$; (b) $d_{(v_2-b_1)}$; (c) $d_{(v_3-b_1)}$; (d) $d_{(v_4-b_1)}$; (e) $d_{(v_5-b_1)}$; and, (f) $d_{(v_6-b_1)}$.




(a) (b)

Figure 4.13: (a) Variable parameter $\frac{K_v}{m}$ and its estimate obtained by MHE and EKF; (b) variable parameter $\frac{K_b}{m}$ and its estimate obtained by MHE and EKF.




(a) (b)

Figure 4.14: (a) boxplot of the RMSEs over 100 runs for the MHE; (b) boxplot of the RMSEs over 100 runs for the EKF.

Figure 4.15 depicts the boxplots of the execution time for the MHE and EKF for one simulation run. From the figure, we can see that MHE has a remarkable larger computational burden as compared to the EKF. Indeed, the average execution time for the EKF to estimate the distances $d_{(v_i-b_1)}$ and the variable parameters $\frac{K_v}{m}$ and $\frac{K_b}{m}$ is 0.2459s, whereas for the MHE the average execution time is 5.9842s which is almost 15 times the time obtained with EKF.

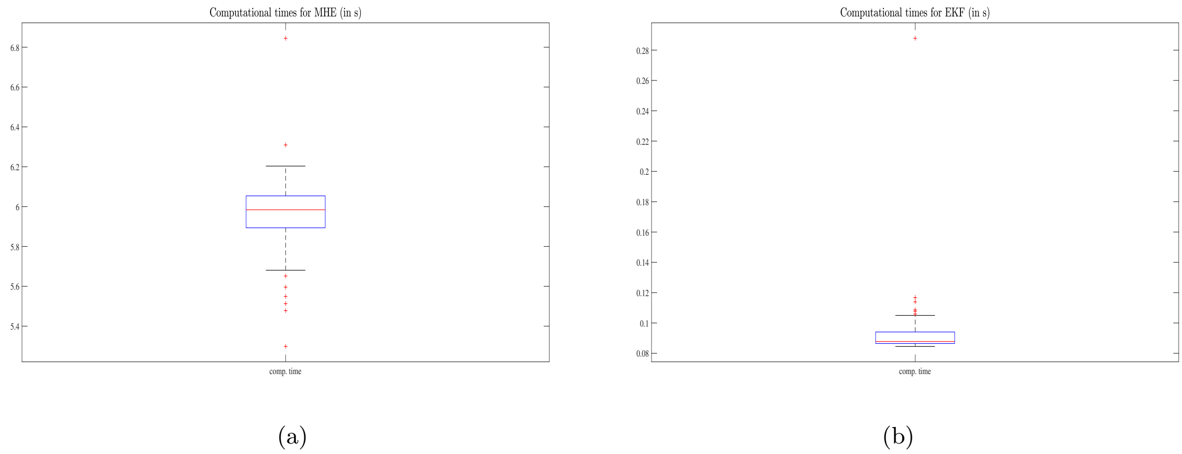



Figure 4.15: (a) Computational time for the MHE; (b) Computational time for the EKF.

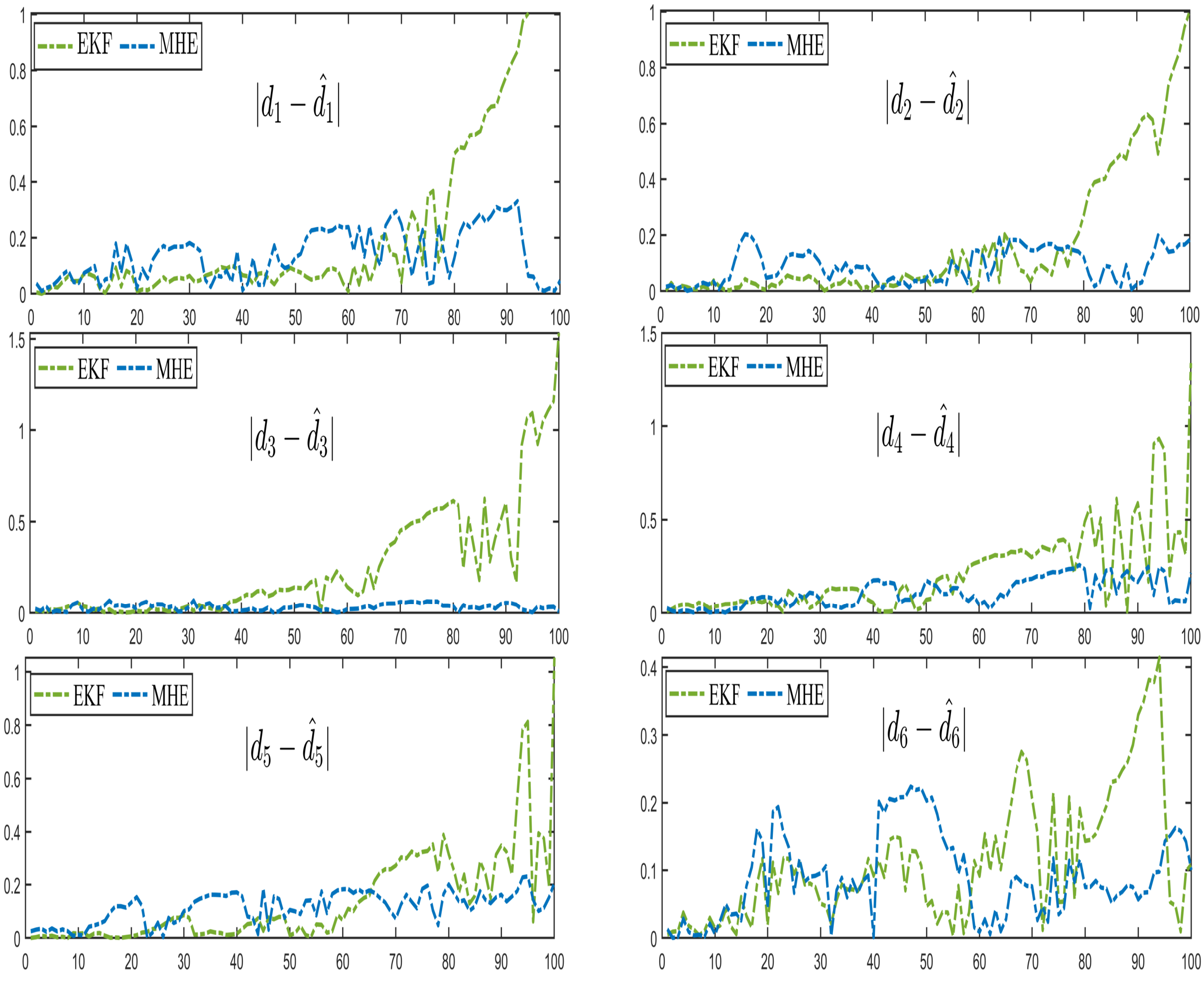



Figure 4.16: Plots of the absolute value of the estimation errors $|d_i - \hat{d}_i|$ where $d_i$ corresponds to the distance between vertex $v_i$ and barycenter $b_1$ and $\hat{d}_i$ is its estimate.

The absolute values of the estimation errors using MHE and EKF are depicted in Figure 4.16, we denote by $d_i$ the distance between vertex $i$ and barycenter $b_1$ and $\hat{d}_i$ its corresponding estimate. According to the obtained plots, it can be concluded that the MHE outperforms the EKF in terms of the stability of the estimation error.

## 4.5 Conclusion

In this chapter, we have investigated three biological models to test the applicability and efficiency of the nonlinear estimation techniques addressed in this thesis, namely, the high-gain observer techniques developed in Chapter 2 and the moving horizon estimation approach presented in Chapter 3.
From the simulation results, we conclude the following points:

- Figures 4.4 and 4.7 from sections (4.2.3) and 4.3.3 respectively, show the remarkable improvements of the new high-gain observer techniques presented in Chapter 2 in terms of peaking phenomenon attenuation, especially for the states $x_2$ and $x_3$. This is due to the smaller value of the newly obtained parameter $\theta_\Gamma$ which is significantly reduced to the power $\frac{1}{(1+j_s)(1+j_0)}$ leading to a smaller value for the observer gain $L_\Gamma$ as compared to the standard high-gain observer $L$.
- The results depicted by Figures 4.5 and 4.8 from sections (4.2.3) and 4.3.3 respectively, clearly put forward the superiority of the new high-gain observer techniques presented in Chapter 2 in terms of sensitivity to measurement noise as compared to the standard high-gain observer. Indeed, this improved observer has the nice feature of reducing the effects of high-frequency measurement noise and eliminating the so-called peaking phenomenon which is typical of the standard high-gain observer.
- The obtained results can be improved by increasing the values of the index parameters $j_s$ and $j_0$, but we can already see from Tables 4.2 and 4.3 that with an index $j_s = 1$ and $j_0 = 1$ the results are quite interesting. In fact, the value of the parameter $\theta$ is significantly reduced from 24.23 with the standard high-gain observer to 2.71 with the improved high-gain observer in the case of the genetic regulatory network and from 43.46 to 3.31 for the SI model.
- In the third and last part of this chapter, we have investigated both modeling and estimation of the Amnioserosa cell mechanics during dorsal closure. The model is described as a quasi-LPV system, and an MHE was considered to estimate the system's states and the unknown parameters jointly. The MHE is accomplished by minimizing a least-squares cost function with respect to both state and parameter variables. The Simulations results depicted in Figures 4.12, 4.16 and 4.13 show the superiority of the MHE as compared to the EKF, indeed, the two estimators are given the same information, namely tuning parameters, model, and measurements; yet MHE consistently provides better precision and robustness to uncertainties compared to the EKF. These benefits of MHE stand for incorporating the physical state constraints into the optimization over a trajectory of states and measurements. Indeed, constraints on the state and on the process and measurement noises can be taken into account directly during the optimization. This can be useful in practice when physical constraints are known
- From the simulation run of the MHE and EKF it is noticed that the MHE is more robust against poor initialization and spurious measurements. This is due to the fact that the estimation is done by optimizing a cost function that takes into account at the same time the initial error, the process noise, and the difference between real measurements and predicted ones. However, the main drawbacks of the MHE are due to its potentially high number of optimization variables. This can lead to strong nonconvexities [177] and the MHE can have difficulties finding the optimal solution. Moreover, large computation time can be induced which would prevent the use of the MHE in real-time applications for systems with fast dynamics.

# Conclusions and perspectives

This thesis presented new results in the field of robust nonlinear state estimation with application to biological plants. In the first part, we contributed to the analytic-based observer through the design of new high-gain estimation techniques to address the issues associated with high-gain observer performance degradation in the presence of measurement noise as well as improving its transient response. The second part is mainly focused on optimization-based state estimation techniques through the moving horizon estimator which is combined with the LPV paradigm, in which we estimate the states and the unknown parameters of a system by explicitly including the unknown parameters in the estimation criterion.

State estimation of linear systems is performed using techniques based on well-established optimal linear estimation theory. Due to the mathematical complexity of nonlinear models, nonlinear state estimation is much less established and requires more analyzing tools. This led us in the first chapter to present a few reminders of the main definitions relating to the stability and observability of nonlinear systems, with much attention given to the notion of observability. The concept of universal inputs and uniform observability is then introduced and a summary of the nonlinear estimation techniques available in the literature closes the first chapter.

The purpose of Chapter 2 is to tackle the challenging performance issues that arise when implementing high-gain observers. A novel methodology for the design of high-gain observers is introduced. It follows the simple methodology of the standard high-gain observer and results in lower values of the observer gain. Indeed, by using the system state augmentation approach, we can augment the dimension of the systems giving rise to a new threshold on the observer parameter $\theta$ which gets smaller as the dimension of the resulting augmented system becomes higher. Furthermore, by combining this design methodology with the HG/LMI technique, a smaller value of $\theta$ is obtained making the observer gain even smaller. This reflects in better transient performances (attenuation of the *"peaking phenomenon"*) which is a typical feature of high-gain observers and less sensitivity to noise in the measurements.

In Chapter 3, the estimation problem is addressed from a different perspective via an optimization-based state estimation approach. The moving horizon estimator is proposed for a class of nonlinear systems that can be written in the form of a quasi-LPV model in the presence of bounded noises affecting both the system and the measurement equations. We have considered the moving-horizon estimator that results from the minimization of a least-squares cost function defined on a sliding window. As compared with the approaches reported in the literature on observers for LPV systems, we have formulated two different approaches denoted "optimistic" and "pessimistic" which allow the inclusion of additional insight into the system's unknown parameters in the form of constraints.

The lack of sensors and reliable procedures to online monitor biological and biomedical processes has drawn our attention in the last chapter at improving and facilitating their operation by applying robust estimation schemes. Towards this end, the nonlinear observation techniques investigated in this thesis are used to estimate the states and parameters of three different models. In the first two sections of Chapter 4, the high-gain estimations techniques presented in Chapter 2 are verified on a simple one-gene regulatory network and a two-stage structured SI epidemic model. From the simulation results, one can note that our proposed design methodologies provide satisfactory results as compared to the standard high-gain observer in terms of sensitivity to high-frequency measurement noise and the peaking phenomenon. In the third section of Chapter 4, we have investigated both modeling and estimation of the Amnioserosa cell mechanics during dorsal closure. The model is described as a quasi-LPV system, and an MHE was considered to jointly estimate the system's states and the unknown parameters. The MHE is accomplished by minimizing a least-squares cost function with respect to both state and parameter variables. Simulation results have been reported to demonstrate the effectiveness of the proposed approach including comparisons with the standard EKF. The two estimators are given the same information, namely tuning parameters, model, and measurements; yet MHE consistently provides better precision and robustness to uncertainties compared to the EKF. These benefits of MHE stand for incorporating the physical state and parameter constraints into the optimization over a trajectory of states and measurements.

Along with this research work, many research opportunities are open for more contributions to the topics encompassed in this thesis. Some of the open problems are presented below :

- Solving observer-based output feedback stabilization problem by using the results on high-gain observers obtained in Chapter 2 is one of the future works we aim to investigate in a deepening way. All the robustness and performance issues related to output feedback control using the state augmentation approach and the HG/LMI observer will be carefully investigated.
- We aim to extend the proposed techniques in Chapter 2 to other classes of nonlinear systems, namely feed-forward systems and systems with time-varying delays in the output measurements. Particular attention will be also given to the problem of real-time adaptation of the observer gain.
- The main advantages of the MHE scheme given in Chapter 3 are its ability to directly incorporate the constraints on the system during the optimization. However, the computation time of the MHE can be very lengthy due to its large number of decision variables. The future main focus involves reducing the online computational complexity to reliably handle challenging large dimensional in nonlinear applications.
- The MHE approach addressed in Chapter 3 uses a standard least-squares type cost function (without an additional maxterm). Robust stability of such an estimator in the presence of bounded disturbances has been observed in practice in many publications but has not been proven so far for general nonlinear detectable systems. However, compared to the case including the additional maxterm, the obtained disturbance gains depend on the estimation horizon $N$ and no uniform bound independent of $N$ could be found. Establishing robust global asymptotic stability of a standard least square type estimator with uniform disturbance gains (i.e., which are valid independent of $N$) is an interesting topic for future research.

Appendix

# A Linear Matrix Inequalities (LMIs)

Linear matrix inequalities (LMIs) were first applied in control system analysis by Yakubovich in the 1960s. Since 1990, the appearance of efficient numerical solvers, such as the ellipsoid method and the interior point algorithm, accelerated the application of LMIs to various engineering problems. Nowadays, it has become a standard numerical tool in control community. Although the LMIs used in some papers may have a very large dimension and look quite complicated, the application of LMIs can make the system analysis and controller design easier.

## A.1 Definition of LMI and BMI

A linear matrix inequality (LMI) is a special semidefinite constraint that has the following form:

$$F(x) = F_0 + x_1 F_1 + \cdots + x_m F_m \leq 0 \tag{A.1}$$

where $x = (x_1, \ldots, x_m)^T$ with $x_i \in \mathbb{R}$ denotes the collection of all the decision variables. $F_i = F_i^T, i = 0, 1, \ldots, m$ are fixed symmetric real matrices. The symbol "$\leq$" represents a negative semidefinite constraint. Similarly, the positive semidefinite constraint can be defined by the symbol " $\geq$ ". From the above definition, the two following properties for $F(x)$ are obvious:

A. Each element in the matrix $F(x)$ is an affine function of $x$;

B. $F(x)$ is a symmetric matrix.

A bilinear matrix inequality is another special semidefinite constraint that has the following form:

$$F(x, y) = F_0 + \sum_{i=1}^{m} x_i F_i + \sum_{j=1}^{n} y_i G_i + \sum_{i=1}^{m} \sum_{j=1}^{n} x_i y_j F_{i,j} \leq 0 \tag{A.2}$$

where $x = (x_1, \ldots, x_m)^T, y = (y_1, \ldots, y_n)^T$ with $x_i, y_j \in \mathbb{R}$ denote the collection of all the decision variables. $F_i = F_i^T, G_i = G_i^T, F_{i,j} = F_{i,j}^T, i = 0, 1, \ldots, m, j = 1, \ldots, n$ are all fixed symmetric real matrices. Similar as the LMI, $F(x, y)$ is a symmetric matrix. Different from the LMI, each element in the matrix $F(x, y)$ is a bi-affine function of $x$ and $y$, which implies that the BMI degenerates to a LMI in $x$ for fixed $y$ and a LMI in $y$ for fixed $x$.

- **Convexity**

The semidefinite condition for $F(x)$ is indeed a convex set for the $x$.

**Lemma A.1**

If the LMI in Eq. (A.1) is feasible, all the feasible points constitute a convex set.

## A.2 Congruence Transformation and Schur Complement

Congruence transformation and Schur complement are two frequently used methods that transfer a BMI constraint to an equivalent LMI constraint. From matrix theory, it is known that the congruence transformation does not change the number of positive, negative and zero eigenvalues of a symmetric matrix. Therefore, an infinite number of equivalent LMIs can be derived.

**Lemma A.2: Congruence Transformation**

The semidefinite condition $F(x) \leq 0$ in Eq. (A.1) or $F(x, y) \leq 0$ is feasible if and only if its congruence transformation is feasible.

$$M^T F(x) M \leq 0, \quad N^T F(x, y) N \leq 0 \tag{A.3}$$

where $M$ and $N$ are both non-singular real matrices.

**Lemma A.3: Schur Complement**

The real symmetric block matrix shown below is negative definite

$$\begin{pmatrix} Q & S^T \\ S & R \end{pmatrix} \prec 0 \tag{A.4}$$

if and only if either of the following two sets of semidefinite constraints is feasible.

A. $Q \prec 0$ and $R - S^T Q^{-1} S \prec 0$

B. $R \prec 0$ and $Q - S R^{-1} S^T \prec 0$

where the symbol $\prec$ represents a negative definite constraint.

## A.3 Feasibility and Optimization

Generally, the application of LMIs in control can be classified into two categories: feasibility and optimization problems.

A. **Feasibility Problem**: This aims at verifying whether there exists at least one point in the decision parameter space such that the given LMI constraint in Eq. (A.1) is feasible. The search for the candidate Lyapunov function for stability analysis is such an example.

B. **Optimization Problem**: This aims at minimizing a linear cost function of $x$ with the LMI in Eq. (A.1) as a constraint.

$$\begin{gathered} \text{minimize}\, c^T x \\ \text{subject to } F(x) \geq 0 \end{gathered}$$

Due to the convexity of the LMI constraint, this is indeed a convex optimization problem. The local minimum is just the global minimum.

Appendix

B

# Lipschitz constant Computation

Consider the nonlinear function $f$

$$f(x) : x \in R^n \to R^n \tag{B.1}$$

which is defined as follows

$$f(x) = \begin{bmatrix} f_1(x)^T & \cdots & f_n(x)^T \end{bmatrix}^T, x = \begin{bmatrix} x_1^T & \cdots & x_n^T \end{bmatrix}^T \tag{B.2}$$

The zero-order Taylor series expansion with integral remainder of $f_i(x)$ around $\hat{x}$ is

$$f_i(x) - f_i(\hat{x}) = \int_{\hat{x}_1}^{x_1} \frac{\partial f_i}{\partial x_1}(t)dt + \ldots + \int_{\hat{x}_n}^{x_n} \frac{\partial f_i}{\partial x_n}(t)dt, \quad i \in \{1, \ldots, n\} \tag{B.3}$$

Every function $f_i$ can be upper bounded as follows

$$|f_i(x) - f_i(\hat{x})| \leq \int_{\hat{x}_1}^{x_1} \left| \frac{\partial f_1}{\partial x_1}(t) \right| dt + \ldots + \int_{\hat{x}_n}^{x_n} \left| \frac{\partial f_1}{\partial x_n}(t) \right| dt \tag{B.4}$$

Given

$$a_{ij} = \max_{t \in [x_j} \hat{x}_j \Big] \left| \frac{\partial f_i}{\partial x_j}(t) \right|, \quad i, j \in \{1, .., n\} \tag{B.5}$$

In the case where the interval $[x_j \quad \hat{x}_j]$ is unknown, then $a_{ij}$ is calculated for $t \in R$; we get

$$a_{ij} = \max_{t \in R} \left| \frac{\partial f_i}{\partial x_j}(t) \right| \tag{B.6}$$

Hence, from (B.4) we can write

$$|f_i(x) - f_i(\hat{x})| \leq a_{i1} |x_1 - \hat{x}_1| + \ldots + a_{in} |x_n - \hat{x}_n| \tag{B.7}$$

By substituting this in $f(x)$, we obtain

$$|f(x) - f(\hat{x})| \leq J|x - \hat{x}| \tag{B.8}$$

where

$$J = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{bmatrix} \tag{B.9}$$

The Lipschitz constant of $f(x)$ is given then by the maximum singular value of $J$.

# Bibliography

## Bibliography


[1] R.E. Kalman. A new approach to linear filtering and prediction problems. *Journal of Basic Engineering*, 82(1):35–45, 1960.

[2] D. Luenberger. An introduction to observers. *IEEE Transactions on Automatic Control*, 16(6):596–602, 1971.

[3] Rudolph E Kalman and Richard S Bucy. New results in linear filtering and prediction theory. 1961.

[4] A.J. Krener and A. Isidori. Linearization by output injection and nonlinear observers. *Systems & Control Letters*, 3:47–52, 1983.

[5] Dieter Bestle and Michael Zeitz. Canonical form observer design for non-linear time-variable systems. *International Journal of Control*, 38:419–431, 1983.

[6] A.J. Krener and W. Respondek. Nonlinear observer with linearizable error dynamics. *SIAM J. on Control and Optimization*, 23(2):197–216, 1985.

[7] X.H. Xia and W.B. Gao. Nonlinear observer design by observer error linearization. *SIAM Journal on Control and Optimization*, 27(1):199–216, 1989.

[8] H Keller. Non-linear observer design by transformation into a generalized observer canonical form. *International Journal of Control*, 46(6):1915–1930, 1987.

[9] Andrew R Phelps. On constructing nonlinear observers. *SIAM Journal on Control and Optimization*, 29(3):516–534, 1991.

[10] J. Rudolph and M. Zeitz. A block triangular nonlinear observer normal form. *Systems & Control Letters*, 23(1):1–8, 1994.

[11] M Hou and AC Pugh. Observer with linear error dynamics for nonlinear multi-output systems. *Systems & Control Letters*, 37(1):1–9, 1999.

[12] Driss Boutat, G Zheng, Jean-Pierre Barbot, and Hassan Hammouri. Observer error linearization multi-output depending. In *Proceedings of the 45th IEEE Conference on Decision and Control*, pages 5394–5399. IEEE, 2006.

[13] Alan F Lynch and Scott A Bortoff. Nonlinear observers with approximately linear error dynamics: The multivariable case. *IEEE Transactions on Automatic Control*, 46(6):927–932, 2001.

[14] MV Lopez, A GLUMIEAU, and F Plestan. Linearization by completely generalized input-output injection. In *Theory and Practice of Control and Systems*, pages 538–543. World Scientific, 1998.

[15] R. Rajamani. Observers for Lipschitz nonlinear systems. *TAC*, 43(3):397–401, 1998.

[16] SEKHAR RAGHAVAN and J. KARL HEDRICK. Observer design for a class of nonlinear systems. *International Journal of Control*, 59(2):515–528, 1994.

[17] Chee Pin Tan and Christopher Edwards. An lmi approach for designing sliding mode observers. *International Journal of Control*, 74(16):1559–1568, 2001.

[18] A. Zemouche and M. Boutayeb. A unified $\mathcal{H}_\infty$ adaptive observer synthesis method for a class of systems with both Lipschitz and monotone nonlinearities. *Systems & Control Letters*, 58(4):282–288, 2009.

[19] Xingzhe Fan and M. Arcak. Nonlinear observer design for systems with multivariable monotone nonlinearities. In *Proceedings of the 41st IEEE Conference on Decision and Control*, volume 1, pages 684–688. IEEE, 2002.

[20] Murat Arcak and Petar KokotoviÄ‡. Nonlinear observers: a circle criterion design and robustness analysis. *Automatica*, 37(12):1923–1930, 2001.

[21] Jean-Jacques E Slotine. Sliding controller design for non-linear systems. *International Journal of control*, 40(2):421–434, 1984.

[22] J-j E Slotine, JK Hedrick, and EA Misawa. Nonlinear state estimation using sliding observers. In *1986 25th IEEE Conference on Decision and Control*, pages 332–339. IEEE, 1986.

[23] BL Walcott and SH Zak. Observation of dynamical systems in the presence of bounded nonlinearities/uncertainties. In *1986 25th IEEE Conference on Decision and Control*, pages 961–966. IEEE, 1986.

[24] B Walcott and S Zak. State observation of nonlinear uncertain dynamical systems. *IEEE Transactions on automatic control*, 32(2):166–170, 1987.

[25] F.E. Thau. Observing the state of non-linear dynamic systems. *Int. J. of Control*, 17(3):471–479, 1973.

[26] J.P. Gauthier, H. Hammouri, and S. Othman. A simple observer for nonlinear systems - application to bioreactors. *IEEE Transactions on Automatic Control*, 37:875–880, 1992.

[27] F. Esfandiari and H. Khalil. Output feedback stabilization of fully linearizable systems. *International Journal of Control*, 56:1007–10037, 1992.

[28] G BesanÃ§on. *Nonlinear Observers and Applications*. Springer, 2007.

[29] K Busawon, M Farza, and H Hammouri. Observer design for a special class of nonlinear systems. *International Journal of Control*, 71(3):405–418, 1998.

[30] F Deza, E Busvelle, JP Gauthier, and D Rakotopara. High gain estimation for nonlinear systems. *Systems & control letters*, 18(4):295–299, 1992.

[31] J.P. Gauthier and I. Kupka. *Deterministic Observation Theory and Applications*. Cambridge University Press, Cambridge, UK, 2001.

[32] J.P. Gauthier and I. Kupka. Observability and observers for nonlinear systems. *SIAM J. on Control and Optimization*, 32(4):975–994, 1994.

[33] H. Hammouri, B. Targui, and F. Armanet. High gain observer based on a triangular structure. *International Journal of Robust and Nonlinear Control*, 12(6):497–518, 2002.

[34] F Viel, Eric Busvelle, and Jean-Paul Gauthier. Stability of polymerization reactors using i/o linearization and a high-gain observer. *Automatica*, 31(7):971–984, 1995.

[35] Hassan K. Khalil. *High-Gain Observers in Nonlinear Feedback Control*. Society for Industrial and Applied Mathematics, Philadelphia, PA, 2017.

[36] Antonio Tornambè. Output feedback stabilization of a class of non-minimum phase nonlinear systems. *Systems & Control Letters*, 19(3):193–204, 1992.

[37] A. Teel and L. Praly. Global stabilizability and observability imply semi-global stabilizability by output feedback. *Systems & Control Letters*, 22(5):313–325, 1994.

[38] A. N. Atassi and H. Khalil. A separation principle for the stabilization of a class of nonlinear systems. *IEEE Transactions on Automatic Control*, 44:1672–1687, 1999.

[39] Hyungbo Shim and Jin Heon Seo. Non-linear output feedback stabilization on a bounded region of attraction. *International Journal of Control*, 73(5):416–426, 2000.

[40] C.I. Byrnes and A. Isidori. Nonlinear internal models for output regulation. *IEEE Transactions on Automatic Control*, 49(12):2244–2247, 2004.

[41] S. Seshagiri and H. Khalil. Robust output feedback regulation of minimum-phase nonlinear systems using conditional integrators. *Automatica*, 41:43–54, 2005.

[42] A.M. Boker and H. Khalil. Nonlinear observers comprising high-gain observers and extended Kalman filters. *Automatica*, 49(12):3583–3590, 2013.

[43] S. Ibrir. Adaptive observers for time-delay nonlinear systems in triangular form. *Automatica*, 45(10):2392–2399, 2009.

[44] J. Ahrens and H. Khalil. High-gain observers in the presence of measurement noise: a switched-gain approach. *Automatica*, 45(4):936–943, 2009.

[45] N. Boizot, E. Busvelle, and J.P Gauthier. An adaptive high-gain observer for nonlinear systems. *Automatica*, 46(9):1483–1488, 2010.

[46] R. Sanfelice and L. Praly. On the performance of high-gain observers with gain adaptation under measurement noise. *Automatica*, 47(10):2165–2176, 2011.

[47] M. Oueder, M. Farza, R. Ben Abdennour, and M. M'Saad. A high gain observer with updated gain for a class of MIMO non-triangular systems. *Systems & Control Letters*, 61(2):298–308, 2012.

[48] V. Andrieu, L. Praly, and A. Astolfi. High gain observers with updated gain and homogeneous correction terms. *Automatica*, 45(2):422–428, 2009.

[49] C. Prieur, S. Tarbouriech, and L. Zaccarian. Hybrid high-gain observers without peaking for planar nonlinear systems. In *51st IEEE Conference on Decision and Control*, pages 6175–6180, Maui, Hawaii, USA, 2012.

[50] A. Alessandri and A. Rossi. Time-varying increasing-gain observers for nonlinear systems. *Automatica*, 49(9):2845–2852, 2013.

[51] A. Alessandri and A. Rossi. Increasing-gain observers for nonlinear systems: stability and design. *Automatica*, 57(7):180–188, 2015.

[52] D. Astolfi and L. Marconi. A high-gain nonlinear observer with limited gain power. *IEEE Transactions on Automatic Control*, 60(11):3059–3064, 2015.

[53] Andrew R Teel. Further variants of the astolfi/marconi high-gain observer. In *2016 american control conference (acc)*, pages 993–998. IEEE, 2016.

[54] H. Khalil. Cascade high-gain observers in output feedback control. *Automatica*, 80:110–118, 2017.

[55] A. Zemouche, F. Zhang, F. Mazenc, and R. Rajamani. High-gain nonlinear observer with lower tuning parameter. *IEEE Transactions on Automatic Control*, 64(8):3194–3209, 2019.

[56] JW Grizzle and PE Moraal. Newton, observers and nonlinear discrete-time control. In *29th IEEE Conference on Decision and Control*, pages 760–767. IEEE, 1990.

[57] H Michalska and DQ Mayne. Moving horizon observers. In *Nonlinear Control Systems Design*, pages 185–190. Elsevier, 1993.

[58] Hannah Michalska and David Q Mayne. Moving horizon observers and observer-based control. *IEEE Transactions on Automatic Control*, 40(6):995–1006, 1995.

[59] PE Moraal and Jessy W Grizzle. Observer design for nonlinear systems with discrete-time measurements. *IEEE Transactions on automatic control*, 40(3):395–404, 1995.

[60] Wolfgang Hahn, Hans H Hosenthien, and H Lehnigk. *Theory and application of Liapunov's direct method*, volume 3. prentice-hall Englewood Cliffs, 1963.

[61] Mathukumalli Vidyasagar. *Nonlinear systems analysis*. SIAM, 2002.

[62] Eduardo D Sontag et al. Smooth stabilization implies coprime factorization. *IEEE transactions on automatic control*, 34(4):435–443, 1989.

[63] SN Dashkovskiy, Denis V Efimov, and Eduardo D Sontag. Input to state stability and allied system properties. *Automation and Remote Control*, 72(8):1579–1614, 2011.

[64] Eduardo D Sontag and Yuan Wang. On characterizations of input-to-state stability with respect to compact sets. In *Nonlinear Control Systems Design 1995*, pages 203–208. Elsevier, 1995.

[65] Eduardo D. Sontag and Yuan Wang. On characterizations of the input-to-state stability property. *Systems & Control Letters*, 24(5):351–359, 1995.

[66] Robert Hermann and Arthur Krener. Nonlinear controllability and observability. *IEEE Transactions on automatic control*, 22(5):728–740, 1977.

[67] J. P. Gauthier and G. Bornard. Observability for any u(t) of a class of nonlinear systems. In *1980 19th IEEE Conference on Decision and Control including the Symposium on Adaptive Processes*, pages 910–915, 1980.

[68] Ali Zemouche. Sur l'observation de l'état des systèmes dynamiques non linéaires. 2007.

[69] Héctor J Sussmann. Single-input observability of continuous-time systems. *Mathematical systems theory*, 12(1):371–393, 1978.

[70] Darrell Williamson. Observation of bilinear systems with application to biological control. *Automatica*, 13(3):243–254, 1977.

[71] Henk Nijmeijer and Arjan J Van der Schaft. *Nonlinear dynamical control systems*, volume 175. Springer, 1990.

[72] Eduardo D Sontag. A concept of local observability. *Systems & Control Letters*, 5(1):41–47, 1984.

[73] Olivier Bernard and J-L Gouzé. State estimation for bioprocesses. Technical report, 2002.

[74] Hendrik Nijmeijer. Observability of autonomous discrete-time non-linear systems: a geometric approach. *International journal of control*, 36(5):867–874, 1982.

[75] BL Walcott, Mí J Corless, and SH Żak. Comparative study of non-linear state-observation techniques. *International Journal of Control*, 45(6):2109–2132, 1987.

[76] Eduardo A Misawa and JK Hedrick. Nonlinear observersâ€”a state-of-the-art survey. 1989.

[77] A. Glumineau, C.H. Moog, and F. Plestan. New algebraic-geometric conditions for the linearization by input-output injection. *IEEE Transactions on Automatic Control*, 41(4):598–603, 1996.

[78] F. Plestan and A. Glumineau. Linearization by generalized input-output injection and synthesis of observers. *IFAC Proceedings Volumes*, 29(1):2162–2166, 1996. 13th World Congress of IFAC, 1996, San Francisco USA, 30 June - 5 July.

[79] Nikolaos Kazantzis and Costas Kravaris. Nonlinear observer design using lyapunovâ€™s auxiliary theorem. *Systems & Control Letters*, 34(5):241–247, 1998.

[80] Ch-W Li and LW Tao. Observing non-linear time-variable systems through a canonical form observer. *International Journal of control*, 44(6):1703–1713, 1986.

[81] M Zeitz. Comments on â€˜comparative study of non-linear state-observation techniquesâ€™. *International Journal of Control*, 46(5):1861–1863, 1987.

[82] Michael Zeitz. The extended luenberger observer for nonlinear systems. *Systems & Control Letters*, 9(2):149–156, 1987.

[83] J Birk and M Zeitz. Extended luenberger observer for non-linear multivariable systems. *International Journal of Control*, 47(6):1823–1836, 1988.

[84] Guanrong Chen. *Approximate kalman filtering*, volume 2. World Scientific, 1993.

[85] Eli Brookner. Tracking and kalman filtering made easy. 1998.

[86] R Brown and P Hwang. *Introduction to Random Signals and Applied Kalman Filtering with MATLAB Exercises*, volume 4. Wiley New York, 2012.

[87] Mohinder S Grewal and Angus P Andrews. *Kalman filtering: Theory and Practice with MATLAB*. John Wiley & Sons, 2014.

[88] Yongkyu Song and Jessy W. Grizzle. The extended kalman filter as a local asymptotic observer for nonlinear discrete-time systems. In *1992 American Control Conference*, pages 3365–3369, 1992.

[89] Andrew H Jazwinski. *Stochastic processes and filtering theory*. Courier Corporation, 2007.

[90] M. Zeitz. The extended luenberger observer for nonlinear systems. *Systems & Control Letters*, 9(2):149–156, 1987.

[91] Murat Arcak. A global separation theorem for a new class of nonlinear observers. In *Proceedings of the 41st IEEE Conference on Decision and Control, 2002.*, volume 1, pages 676–681. IEEE, 2002.

[92] Murat Arcak and Petar Kokotovic. Observer-based control of systems with slope-restricted nonlinearities. *IEEE Transactions on Automatic Control*, 46(7):1146–1150, 2001.

[93] Xingzhe Fan and Murat Arcak. Observer design for systems with multivariable monotone nonlinearities. *Systems & Control Letters*, 50(4):319–330, 2003.

[94] J. Gauthier and G. Bornard. Observability for anyu(t)of a class of nonlinear systems. *IEEE Transactions on Automatic Control*, 26(4):922–926, 1981.

[95] M. Zeitz. Observability canonical (phase-variable) form for nonlinear time-variable systems. *International Journal of Systems Science*, 15(9):949â€“958„ 1984.

[96] A. Tornambe. Use of asymptotic observers having-high-gains in the state and parameter estimation. In *Proceedings of the 28th IEEE Conference on Decision and Control,*, pages 1791–1794 vol.2, 1989.

[97] G. Bornard and H. Hammouri. A high gain observer for a class of uniformly observable systems. In *[1991] Proceedings of the 30th IEEE Conference on Decision and Control*, pages 1494–1496 vol.2, 1991.

[98] Shauying R Kou, David L Elliott, and Tzyh Jong Tarn. Exponential observers for nonlinear dynamic systems. *Information and control*, 29(3):204–216, 1975.

[99] G Ciccarella, M Dalla Mora, and Alfredo Germani. A luenberger-like observer for nonlinear systems. *International Journal of Control*, 57(3):537–556, 1993.

[100] Gerhard Kreisselmeier and Robert Engel. Nonlinear observers for autonomous lipschitz continuous systems. *IEEE Transactions on Automatic Control*, 48(3):451–464, 2003.

[101] Fanglai Zhu and Zhengzhi Han. A note on observers for lipschitz nonlinear systems. *IEEE Transactions on automatic control*, 47(10):1751–1754, 2002.

[102] G Bornard, F Celle-Couenne, and G Gilles. Observability and observers. In *Nonlinear systems*, pages 173–216. Springer, 1995.

[103] Angelo Alessandri. Observer design for nonlinear systems by using input-to-state stability. In *2004 43rd IEEE Conference on Decision and Control (CDC)(IEEE Cat. No. 04CH37601)*, volume 4, pages 3892–3897. IEEE, 2004.

[104] Mondher Farza, Krishna Busawon, and Houbekeur Hammouri. Simple nonlinear observers for on-line estimation of kinetic rates in bioreactors. *Automatica*, 34(3):301–318, 1998.

[105] Olivier Bernard, Gauthier Sallet, and Antoine Sciandra. Nonlinear observers for a class of biological systems: application to validation of a phytoplanktonic growth model. *IEEE Transactions on Automatic Control*, 43(8):1056–1065, 1998.

[106] J-JE Slotine, J Kand Hedrick, and Eduardo A Misawa. On sliding observers for nonlinear systems. *Journal of Dynamic Systems*, 109:245â€“–252, 1987.

[107] B Walcott and S Zak. State observation of nonlinear uncertain dynamical systems. *IEEE Transactions on automatic control*, 32(2):166–170, 1987.

[108] DM Dawson, Z Qu, and JC Carroll. On the state observation and output feedback problems for nonlinear uncertain dynamic systems. *Systems & Control Letters*, 18(3):217–222, 1992.

[109] Kenneth R Muske and James B Rawlings. Nonlinear moving horizon state estimation. In *Methods of model based process control*, pages 349–365. Springer, 1995.

[110] Douglas G Robertson, Jay H Lee, and James B Rawlings. A moving horizon-based approach for least-squares estimation. *AIChE Journal*, 42(8):2209–2224, 1996.

[111] Rawlings James Blake and Mayne David. *Model Predictive Control: Theory and Design.* Nob Hill Publishing, 2009.

[112] Christopher V Rao. *Moving horizon strategies for the constrained monitoring and control of nonlinear discrete-time systems.* The University of Wisconsin-Madison, 2000.

[113] Peter Klaus Findeisen. *Moving horizon state estimation of discrete time systems.* PhD thesis, University of Wisconsin–Madison, 1997.

[114] Matthew Jeremy Tenny. *Computational strategies for nonlinear model predictive control.* The University of Wisconsin-Madison, 2002.

[115] Angelo Alessandri, Marco Baglietto, and Giorgio Battistelli. Receding-horizon estimation for discrete-time linear systems. *IEEE Transactions on Automatic Control*, 48(3):473–478, 2003.

[116] Mark L Darby and Michael Nikolaou. A parametric programming approach to moving-horizon state estimation. *Automatica*, 43(5):885–891, 2007.

[117] H. Khalil. *Nonlinear Systems.* Prentice Hall, Upper Saddle River, NJ, 2002.

[118] Hassan K Khalil and Laurent Praly. High-gain observers in nonlinear feedback control. *International Journal of Robust and Nonlinear Control*, 24(6):993–1015, 2014.

[119] W.M. Wonham. *Linear Multivariable Control - A geometric approach.* Springer Verlag, New York, 1985.

[120] T Mita. On zeros and responses of linear regulators and linear observers. *IEEE Transactions on Automatic Control*, 22(3):423–428, 1977.

[121] M. Maggiore and K.M. Passino. A separation principle for a class of non-uco systems. *IEEE Transactions on Automatic Control*, 48(7):1122–1133, 2003.

[122] E.H. El Yaagoubi, A. El Assoudi, and H. Hammouri. High gain observer: attenuation of the peak phenomena. In *Proceedings of the 2004 American Control Conference*, volume 5, pages 4393–4397 vol.5, 2004.

[123] Alexis A Prasov and Hassan K Khalil. A nonlinear high-gain observer for systems with measurement noise in a feedback control framework. *IEEE Transactions on Automatic Control*, 58(3):569–580, 2012.

[124] A. Zemouche and M. Boutayeb. On LMI conditions to design observers for Lipschitz nonlinear systems. *Automatica*, 49(2):585–591, 2013.

[125] A Jazwinski. Limited memory optimal filtering. *IEEE Transactions on Automatic Control*, 13(5):558–563, 1968.

[126] C.V. Rao, J.B. Rawlings, and J.H. Lee. Constrained linear estimation – a moving horizon approach. *Automatica*, 37(10):1619–1628, 2001.

[127] Eric L Haseltine and James B Rawlings. Critical evaluation of extended kalman filtering and moving-horizon estimation. *Industrial & engineering chemistry research*, 44(8):2451–2460, 2005.

[128] Qinyuan Liu and Zidong Wang. Moving-horizon estimation for linear dynamic networks with binary encoding schemes. *IEEE Transactions on Automatic Control*, 66(4):1763–1770, 2020.

[129] Shi Shang Jang, Babu Joseph, and Hiro Mukai. Comparison of two approaches to on-line parameter and state estimation of nonlinear systems. *Industrial & Engineering Chemistry Process Design and Development*, 25(3):809–814, 1986.

[130] G. Zimmer. State observation by on-line minimization. *International Journal of Control*, 60:595–606, 1994.

[131] Angelo Alessandri, Marco Baglietto, and Giorgio Battistelli. On estimation error bounds for receding-horizon filters using quadratic boundedness. *IEEE Transactions on Automatic Control*, 49(8):1350–1355, 2004.

[132] A. Alessandri, M. Baglietto, and G. Battistelli. Robust receding-horizon state estimation for uncertain discrete-time linear systems. *Systems & Control Letters*, 54(7):627–643, 2005.

[133] G. Ferrari-Trecate, D. Mignone, and M. Morari. Moving horizon estimation for hybrid systems. *IEEE Transactions on Automatic Control*, 47(10):1663–1676, 2002.

[134] Angelo Alessandri, Marco Baglietto, and Giorgio Battistelli. Receding-horizon estimation for switching discrete-time linear systems. *IEEE Transactions on Automatic Control*, 50(11):1736–1748, 2005.

[135] Mazen Alamir. Optimization-based non-linear observers revisited. *International Journal of Control*, 72(13):1204–1217, 1999.

[136] C.V. Rao, J.B. Rawlings, and D.Q. Mayne. Constrained state estimation for nonlinear discrete-time systems: stability and moving horizon approximations. *IEEE Trans. on Automatic Control*, 48(2):246–257, 2003.

[137] A. Alessandri, M. Baglietto, T. Parisini, and R. Zoppoli. A neural state estimator with bounded errors for nonlinear systems. *IEEE Transactions on Automatic Control*, 44(11):2028–2042, 1999.

[138] A. Alessandri, M. Baglietto, and G. Battistelli. Min-max moving-horizon estimation for uncertain discrete-time systems. *SIAM Journal on Control and Optimization*, 50(3):1439–1465, 2012.

[139] Marcello Farina, Giancarlo Ferrari-Trecate, and Riccardo Scattolini. Distributed moving horizon estimation for linear constrained systems. *IEEE Transactions on Automatic Control*, 55(11):2462–2475, 2010.

[140] René Schneider, Ralf Hannemann-Tamás, and Wolfgang Marquardt. An iterative partition-based moving horizon estimator with coupled inequality constraints. *Automatica*, 61:302–307, 2015.

[141] A Ronald Gallant. *Nonlinear statistical models*. John Wiley & Sons, 2009.

[142] G. Bastin and D. Dochain. On-line estimation of microbial specific growth rates. *Automatica*, 22(6):705–709, 1986.

[143] Georges Bastin. *On-line estimation and adaptive control of bioreactors*, volume 1. Elsevier, 2013.

[144] Jean-Jacques E Slotine, Weiping Li, et al. *Applied nonlinear control*, volume 199. Prentice hall Englewood Cliffs, NJ, 1991.

[145] D Bonvin, P De Vallie, DWT Rippin, et al. Application of estimation techniques to batch reactorsâ€"i. modelling thermal effects. *Computers & chemical engineering*, 13(1-2):1–9, 1989.

[146] Ph De Valliere and D Bonvin. Application of estimation techniques to batch reactorsâ€"iii. modelling refinements which improve the quality of state and parameter estimation. *Computers & chemical engineering*, 14(7):799–808, 1990.

[147] RD Gudi, SL Shah, and MR Gray. Multirate state and parameter estimation in an antibiotic fermentation with delayed measurements. *Biotechnology and Bioengineering*, 44(11):1271–1278, 1994.

[148] D. Semino, M. Morretta, and C. Scali. Parameter estimation in extended kalman filters for quality control in polymerization reactors. *Computers & Chemical Engineering*, 20:S913–S918, 1996. European Symposium on Computer Aided Process Engineering-6.

[149] Ashuraj Sirohi and Kyu Yong Choi. On-line parameter estimation in a continuous polymerization process. *Industrial & engineering chemistry research*, 35(4):1332–1343, 1996.

[150] Masoud Soroush. Nonlinear state-observer design with application to reactors. *Chemical Engineering Science*, 52(3):387–404, 1997.

[151] Long-Wen Liu, Wei Xie, and Lang-Wen Zhang. Reduced-order interval observer design for continuous-time descriptor lpv systems with uncertainties. *International Journal of Control*, pages 1–12, 2022.

[152] Dalil Ichalal, Said Mammar, and José Ragot. Auxiliary dynamics for observer design of nonlinear ts systems with unmeasurable premise variables. *IFAC-PapersOnLine*, 49(5):1–6, 2016.

[153] Habib Hamdi, Mickael Rodrigues, Chokri Mechmeche, Didier Theilliol, and N BenHadj Braiek. Fault detection and isolation in linear parameter-varying descriptor systems via proportional integral observer. *International journal of adaptive control and signal processing*, 26(3):224–240, 2012.

[154] Denis Efimov, Tarek Raïssi, and Ali Zolghadri. Control of nonlinear and lpv systems: interval observer-based framework. *IEEE Transactions on Automatic Control*, 58(3):773–778, 2013.

[155] Yan Wang, David M Bevly, and Rajesh Rajamani. Interval observer design for lpv systems with parametric uncertainty. *Automatica*, 60:79–85, 2015.

[156] Stanislav Chebotarev, Denis Efimov, Tarek Raïssi, and Ali Zolghadri. Interval observers for continuous-time lpv systems with l1/l2 performance. *Automatica*, 58:82–89, 2015.

[157] Stefan Krebs, Martin Pfeifer, Sebastian Fugel, Jörg Weigold, and Sören Hohmann. Interval observer for lpv systems based on time-variant transformations. In *2016 IEEE 55th Conference on Decision and Control (CDC)*, pages 4090–4096. IEEE, 2016.

[158] Nan Gao, Mohamed Darouach, Holger Voos, and Marouane Alma. New unified hâˆž dynamic observer design for linear systems with unknown inputs. *Automatica*, 65:43–52, 2016.

[159] A-J Pérez-Estrada, G-L Osorio-Gordillo, M Darouach, M Alma, and V-H Olivares-Peregrino. Generalized dynamic observers for quasi-lpv systems with unmeasurable scheduling functions. *International Journal of Robust and Nonlinear Control*, 28(17):5262–5278, 2018.

[160] Souad Bezzaoucha, Holger Voos, and Mohamed Darouach. A new polytopic approach for the unknown input functional observer design. *International Journal of Control*, 91(3):658–677, 2018.

[161] A-J Pérez-Estrada, G-L Osorio-Gordillo, Mohamed Darouach, and Marouane Alma. Adaptive observer design for lpv systems. *IFAC-PapersOnLine*, 52(28):140–145, 2019.

[162] Denis Efimov, Tarek Raissi, and Ali Zolghadri. Set adaptive observers for linear parameter-varying systems: Application to fault detection. *Journal of Dynamic Systems, Measurement, and Control*, 136(2):7, 2014.

[163] Mazen Alamir. Nonlinear moving horizon observers: Theory and real-time implementation. In *nonlinear observers and applications*, pages 139–179. Springer, 2007.

[164] Torsten Söderström. *Discrete-time stochastic systems: estimation and control.* Springer Science & Business Media, 2002.

[165] Arthur Gelb et al. *Applied optimal estimation.* MIT press, 1974.

[166] Arthur E Bryson and Yu-Chi Ho. *Applied optimal control: optimization, estimation, and control.* Routledge, 2018.

[167] Robert F Stengel. *Optimal control and estimation.* Courier Corporation, 1994.

[168] S Sathiya Keerthi and Elmer G Gilbert. Optimal infinite-horizon feedback laws for a general class of constrained discrete-time systems: Stability and moving-horizon approximations. *Journal of optimization theory and applications*, 57(2):265–293, 1988.

[169] Kenneth R Muske, James B Rawlings, and Jay H Lee. Receding horizon recursive state estimation. In *1993 American Control Conference*, pages 900–904. IEEE, 1993.

[170] Y Ramamurthi, PB Sistu, and BW Bequette. Control-relevant dynamic data reconciliation and parameter estimation. *Computers & chemical engineering*, 17(1):41–59, 1993.

[171] I-W Kim, MJ Liebman, and TF Edgar. A sequential error-in-variables method for nonlinear dynamic systems. *Computers & chemical engineering*, 15(9):663–670, 1991.

[172] S. T. Glad. Observability and nonlinear dead beat observers. In *The 22nd IEEE Conference on Decision and Control*, pages 800–802, 1983.

[173] Angelo Alessandri, Marco Baglietto, Thomas Parisini, and Riccardo Zoppoli. A neural state estimator with bounded errors for nonlinear systems. *IEEE Transactions on Automatic Control*, 44(11):2028–2042, 1999.

[174] A. Alessandri, M. Baglietto, and G. Battistelli. Moving-horizon state estimation for nonlinear discrete-time systems: New stability results and approximation schemes. *Automatica*, 44(7):1753–1765, 2008.

[175] Tobias Raff, Christian Ebenbauer, Rolf Findeisen, and Frank Allgöwer. Remarks on moving horizon state estimation with guaranteed convergence. In *Control and Observer Design for Nonlinear Finite and Infinite Dimensional Systems*, pages 67–80. Springer, 2005.

[176] Emrah BÄ±yÄ±k and Murat Arcak. A hybrid redesign of newton observers in the absence of an exact discrete-time model. *Systems & Control Letters*, 55(6):429–436, 2006.

[177] Victor M Zavala, Carl D Laird, and Lorenz T Biegler. A fast moving horizon estimation algorithm based on nonlinear programming sensitivity. *Journal of Process Control*, 18(9):876–884, 2008.

[178] Dan Sui, Tor Arne Johansen, and Le Feng. Linear moving horizon estimation with pre-estimating observer. *IEEE Transactions on automatic control*, 55(10):2363–2368, 2010.

[179] H. Cox. On the estimation of state variables and parameters for noisy dynamic systems. *IEEE Transactions on Automatic Control*, 9(1):5–12, 1964.

[180] Lennart Ljung. Asymptotic behavior of the extended kalman filter as a parameter estimator for linear systems. *IEEE Transactions on Automatic Control*, 24(1):36–50, 1979.

[181] Ph De Valliere and D Bonvin. Application of estimation techniques to batch reactorsâ€”iii. modelling refinements which improve the quality of state and parameter estimation. *Computers & chemical engineering*, 14(7):799–808, 1990.

[182] M. Agarwal and D. Bonvin. Limitations of the extended kalman filter for batch reactors. *IFAC Proceedings Volumes*, 22(8):299–306, 1989. IFAC Symposium on Dynamics and Control of Chemical Reactors, Distillation Columns and Batch Processes , Maastricht, The Netherlands, 21-23 August 1989.

[183] Wilson J Rugh and Jeff S Shamma. Research on gain scheduling. *Automatica*, 36(10):1401–1425, 2000.

[184] Gary J Balas. Linear, parameter-varying control and its application to a turbofan engine. *International Journal of Robust and Nonlinear Control: IFAC-Affiliated Journal*, 12(9):763–796, 2002.

[185] Carsten W. Scherer and Siep Weiland. Lecture notes disc course on linear matrix inequalities in control. 1999.

[186] Angelo Alessandri, Marco Baglietto, and Giorgio Battistelli. Moving-horizon state estimation for nonlinear discrete-time systems: New stability results and approximation schemes. *Automatica*, 44(7):1753–1765, 2008.

[187] Ali Zemouche, Rajesh Rajamani, Houria Kheloufi, and Fazia Bedouhene. Robust observer-based stabilization of lipschitz nonlinear uncertain systems via lmis-discussions and new design procedure. *International journal of robust and nonlinear control*, 27(11):1915–1939, 2017.

[188] Thierry Marie Guerra, Victor Estrada-Manzo, and Zsófia Lendek. Observer design for takagi–sugeno descriptor models: An lmi approach. *Automatica*, 52:154–159, 2015.

[189] D. Ichalal, B. Marx, S. Mammar, D. Maquin, and J. Ragot. How to cope with unmeasurable premise variables in takagiâ€“sugeno observer design: Dynamic extension approach. *Engineering Applications of Artificial Intelligence*, 67:430–435, 2018.

[190] Ali Zemouche, Rajesh Rajamani, Gridsada Phanomchoeng, Boulaïd Boulkroune, Hugues Rafaralahy, and Michel Zasadzinski. Circle criterion-based â„‹âˆž observer design for lipschitz and monotonic nonlinear systems–enhanced lmi conditions and constructive discussions. *Automatica*, 85:412–425, 2017.

[191] Young Man Cho and Rajesh Rajamani. A systematic approach to adaptive observer synthesis for nonlinear systems. *IEEE transactions on Automatic Control*, 42(4):534–537, 1997.

[192] Liyi Dai. *Singular control systems*. Springer, 1989.

[193] M. Dureau, A. Alessandri, P. Bagnerini, and S. Vincent. Modeling and identification of amnioserosa cell mechanical behavior by using mass-spring lattices. *IEEE/ACM Trans. on Computational Biology and Bioinformatics*, 14(6):1476â€“1481, 2017.

[194] B. Lewin. *Genes VII*. Oxford University Press, 1999.

[195] L.A Torres, V. Ibarra-Junquerab, P. Escalante-Minakataa, and H.C. Rosua. High-gain nonlinear observer for simple genetic regulation process. *Physica A*, 380:235â€“–240, 2007.

[196] David J Lockhart and Elizabeth A Winzeler. Genomics, gene expression and dna arrays. *Nature*, 405(6788):827–836, 2000.

[197] Mark Isalan, Caroline Lemerle, Konstantinos Michalodimitrakis, Carsten Horn, Pedro Beltrao, Emanuele Raineri, Mireia Garriga-Canut, and Luis Serrano. Evolvability and hierarchy in rewired bacterial gene networks. *Nature*, 452(7189):840–845, 2008.

[198] Trey Ideker, Vesteinn Thorsson, Jeffrey A Ranish, Rowan Christmas, Jeremy Buhler, Jimmy K Eng, Roger Bumgarner, David R Goodlett, Ruedi Aebersold, and Leroy Hood. Integrated genomic and proteomic analyses of a systematically perturbed metabolic network. *Science*, 292(5518):929–934, 2001.

[199] Sourav Bandyopadhyay, Monika Mehta, Dwight Kuo, Min-Kyung Sung, Ryan Chuang, Eric J Jaehnig, Bernd Bodenmiller, Katherine Licon, Wilbert Copeland, Michael Shales, et al. Rewiring of genetic networks in response to dna damage. *Science*, 330(6009):1385–1389, 2010.

[200] Han-Yu Chuang, Eunjung Lee, Yu-Tsueng Liu, Doheon Lee, and Trey Ideker. Network-based classification of breast cancer metastasis. *Molecular systems biology*, 3(1):140, 2007.

[201] Zidong Wang, James Lam, Guoliang Wei, Karl Fraser, and Xiaohui Liu. Filtering for nonlinear genetic regulatory networks with stochastic disturbances. *IEEE Transactions on Automatic Control*, 53(10):2448–2457, 2008.

[202] Pagavathigounder Balasubramaniam and L Jarina Banu. Robust state estimation for discrete-time genetic regulatory network with random delays. *Neurocomputing*, 122:349–369, 2013.

[203] R Sakthivel, K Mathiyalagan, S Lakshmanan, and Ju H Park. Robust state estimation for discrete-time genetic regulatory networks with randomly occurring uncertainties. *Nonlinear dynamics*, 74(4):1297–1315, 2013.

[204] Tong Wang, Yongsheng Ding, Lei Zhang, and Kuangrong Hao. Robust state estimation for discrete-time stochastic genetic regulatory networks with probabilistic measurement delays. *Neurocomputing*, 111:1–12, 2013.

[205] Dan Zhang, Haiyu Song, Li Yu, Qing-Guo Wang, and Chongjin Ong. Set-values filtering for discrete time-delay genetic regulatory networks with time-varying parameters. *Nonlinear Dynamics*, 69(1-2):693–703, 2012.

[206] Xian Zhang, Yuanyuan Han, Ligang Wu, and Yantao Wang. State estimation for delayed genetic regulatory networks with reaction–diffusion terms. *IEEE Transactions on neural networks and learning systems*, 29(2):299–309, 2016.

[207] Ahmet Ay and David N Arnosti. Mathematical modeling of gene expression: a guide for the perplexed biologist. *Critical reviews in biochemistry and molecular biology*, 46(2):137–151, 2011.

[208] J.J. Tyson and H.G. Othmer. The dynamics of feedback control circuits in biochemical pathways. *Progress in theoretical biology*, 5:1â€“–62, 1978.

[209] B.C. Goodwin. *Temporal organization in cells*. Academic Press, 1963.

[210] B.C. Goodwin. Oscillatory behavior in enzymatic control processes. *Advances in Enzyme Regulation*, 3:425–428, 1965.

[211] World Health Organization. Hiv/aids. global health observatory (gho) data.

[212] Laetitia Vaillant, Guy La Ruche, Arnaud Tarantola, Philippe Barboza, et al. Epidemiology of fatal cases associated with pandemic h1n1 influenza 2009. *Eurosurveillance*, 14(33):1–6, 2009.

[213] WHO Ebola Response Team. Ebola virus disease in west africaâ€”the first 9 months of the epidemic and forward projections. *New England Journal of Medicine*, 371(16):1481–1495, 2014.

[214] Adam J Kucharski, Timothy W Russell, Charlie Diamond, Yang Liu, John Edmunds, Sebastian Funk, Rosalind M Eggo, Fiona Sun, Mark Jit, James D Munday, et al. Early dynamics of transmission and control of covid-19: a mathematical modelling study. *The lancet infectious diseases*, 20(5):553–558, 2020.

[215] Matthias Egger, Leigh Johnson, Christian Althaus, Anna Schöni, Georgia Salanti, Nicola Low, and Susan L Norris. Developing who guidelines: time to formally include evidence from mathematical modelling studies. *F1000Research*, 6, 2017.

[216] Nicholas P Jewell, Joseph A Lewnard, and Britta L Jewell. Predictive mathematical models of the covid-19 pandemic: underlying principles and value of projections. *Jama*, 323(19):1893–1894, 2020.

[217] David Adam. Special report: The simulations driving the world's response to covid-19. *Nature*, 580(7802):316–319, 2020.

[218] Dongmei Chen, Bernard Moulin, and Jianhong Wu. *Analyzing and modeling spatial and temporal dynamics of infectious diseases*. John Wiley & Sons, 2014.

[219] Sheryl L Chang, Nathan Harding, Cameron Zachreson, Oliver M Cliff, and Mikhail Prokopenko. Modelling transmission and control of the covid-19 pandemic in australia. *Nature communications*, 11(1):1–13, 2020.

[220] Reza Sameni. Mathematical modeling of epidemic diseases; a case study of the covid-19 coronavirus. *arXiv preprint arXiv:2003.11371*, 2020.

[221] William Ogilvy Kermack and Anderson G McKendrick. A contribution to the mathematical theory of epidemics. *Proceedings of the royal society of london. Series A, Containing papers of a mathematical and physical character*, 115(772):700–721, 1927.

[222] Abderrahman Iggidr and Max O Souza. State estimators for some epidemiological systems. *Journal of mathematical biology*, 78(1):225–256, 2019.

[223] Stefania Audoly, Leontina D'Angio, Maria Pia Saccomani, and Claudio Cobelli. Global identifiability of linear compartmental models-a computer algebra algorithm. *IEEE Transactions on Biomedical Engineering*, 45(1):36–47, 1998.

[224] John A Jacquez and Peter Greif. Numerical parameter identifiability and estimability: Integrating identifiability, estimability, and optimal sampling design. *Mathematical Biosciences*, 77(1-2):201–227, 1985.

[225] Tina Toni, David Welch, Natalja Strelkowa, Andreas Ipsen, and Michael PH Stumpf. Approximate bayesian computation scheme for parameter inference and model selection in dynamical systems. *Journal of the Royal Society Interface*, 6(31):187–202, 2009.

[226] Abdessamad Abdelhedi, Driss Boutat, Lassaad Sbita, and Ramdane Tami. Extended observer to estimate the spreading of contagious disease. In *2014 European Control Conference (ECC)*, pages 1843–1848, 2014.

[227] Jose Alvarez-Ramirez, Monica Meraz, and Jorge X Velasco-Hernandez. Feedback control of the chemotherapy of hiv. *International Journal of Bifurcation and Chaos*, 10(09):2207–2219, 2000.

[228] Derdei Bichara, Nathalie Cozic, and Abderrahman Iggidr. On the estimation of sequestered infected erythrocytes in plasmodium falciparum malaria patients. *Mathematical Biosciences and engineering*, 11(4):741–759, 2014.

[229] Ramdane Tami, Driss Boutat, Gang Zheng, and Frederic Kratz. Parameters and states estimation for dengue epidemic model. In *2014 European Control Conference (ECC)*, pages 528–533, 2014.

[230] Asier Ibeas, Manuel de la Sen, Santiago Alonso-Quesada, and Iman Zamani. Stability analysis and observer design for discrete-time seir epidemic models. *Advances in Difference Equations*, 2015(1):1–21, 2015.

[231] Linda J.S. Allen. Some discrete-time si, sir, and sis epidemic models. *Mathematical Biosciences*, 124(1):83–105, 1994.

[232] Robert Smith. *Modelling Disease Ecology with Mathematics*. American Institute of Mathematical Sciences, 2008.

[233] Mamadou L Diouf, Abderrahman Iggidr, and Max O Souza. Stability and estimation problems related to a stage-structured epidemic model. *Mathematical Biosciences and Engineering*, 16(5):4415–4432, 2019.

[234] Alexander G Fletcher, James M Osborne, Philip K Maini, and David J Gavaghan. Implementing vertex dynamics models of cell populations in biology within a consistent computational framework. *Progress in biophysics and molecular biology*, 113(2):299–326, 2013.

[235] D.P. Kiehart. Wound healing: The power of the purse string. *Current Biology*, 9(16):602â€“605, 1999.

[236] Daniel P Kiehart, Catherine G. Galbraith, Kevin A. Edwards, Wayne L. Rickoll, and Ruth A. Montague. Multiple forces contribute to cell sheet morphogenesis for dorsal closure in drosophila. *The Journal of Cell Biology*, 149:471 – 490, 2000.

[237] Antonio Jacinto, Sarah Woolner, and Paul Martin. Dynamic analysis of dorsal closure in drosophila: from genetics to cell biology. *Developmental cell*, 3(1):9–19, 2002.

[238] Nicholas Harden. Signaling pathways directing the movement and fusion of epithelial sheets: lessons from dorsal closure in drosophila. *Differentiation; research in biological diversity*, 70 4-5:181–203, 2002.

[239] Antonio Jacinto and Paul Martin. Morphogenesis: unravelling the cell biology of hole closure. *Current Biology*, 11(17):R705–R707, 2001.

[240] Nicole Gorfinkiel, Guy B Blanchard, Richard J Adams, and Alfonso Martinez Arias. Mechanical control of global cell behaviour during dorsal closure in drosophila. 2009.

[241] Nicole Gorfinkiel and Alfonso Martinez Arias. Requirements for adherens junction components in the interaction between epithelial tissues during dorsal closure in drosophila. *Journal of Cell Science*, 120:3289 – 3298, 2007.

[242] Josef D. Franke, Ruth A. Montague, and Daniel P Kiehart. Nonmuscle myosin ii generates forces that transmit tension and drive contraction in multiple tissues during dorsal closure. *Current Biology*, 15:2208–2221, 2005.

[243] Michele L Lamka and Howard D Lipshitz. Role of the amnioserosa in germ band retraction of the drosophila melanogaster embryo. *Developmental biology*, 214(1):102–112, 1999.

[244] Anne Scuderi and Anthea Letsou. Amnioserosa is required for dorsal closure in drosophila. *Developmental Dynamics*, 232(3):791–800, 2005.

[245] AC Aristotelous, JM Crawford, GS Edwards, DP Kiehart, and S Venakides. Mathematical models of dorsal closure. *Progress in biophysics and molecular biology*, 137:111–131, 2018.

[246] Hiroaki Kojima, Akihiko Ishijima, and Toshio Yanagida. Direct measurement of stiffness of single actin filaments with and without tropomyosin by in vitro nanomanipulation. *Proceedings of the National Academy of Sciences*, 91(26):12962–12966, 1994.

[247] Md. Mozzammel Haque, Mihaela G. Moisescu, Sándor Valkai, András Dér, and Tudor Savopol. Stretching of red blood cells using an electro-optics trap. *Biomed. Opt. Express*, 6(1):118–123, 2015.

[248] M. Murrell and ML. Gardel. Actomyosin sliding is attenuated in contractile biomimetic cortices. *Mol Biol Cell*, 25(12):1845–1853, 2014.

[249] Karen E. Kasza, Amy C. Rowat, Jiayu Liu, Thomas E. Angelini, Clifford P. Brangwynne, Gijsje Hendrika Koenderink, and David A. Weitz. The cell as a material. *Current opinion in cell biology*, 19 1:101–7, 2007.

[250] M.S. Hutson, J. Veldhuis, X. Ma, H.E. Lynch, P.G. Cranston, and G.W. Brodland. Combining laser microsurgery and finite element modeling to assess cell-level epithelial mechanics. *Biophysical Journal*, 97(12):3075â€“3085, 2009.

[251] S Ishihara, K Sugimura, SJ Cox, Isabelle Bonnet, Y Bellaïche, and F Graner. Comparative study of non-invasive force and stress inference methods in tissue. *The European Physical Journal E*, 36(4):1–13, 2013.

[252] G. Wayne Brodland, Vito Conte, P. Graham Cranston, Jim Veldhuis, Sriram Narasimhan, M. Shane Hutson, Antonio Jacinto, Florian Ulrich, Buzz Baum, Mark Miodownik, and William A. Eaton. Video force microscopy reveals the mechanics of ventral furrow invagination in drosophila. *Proceedings of the National Academy of Sciences of the United States of America*, 107(51):22111–22116, 2010.

[253] Guy B Blanchard, Sughashini Murugesu, Richard J Adams, Alfonso Martinez-Arias, and Nicole Gorfinkiel. Cytoskeletal dynamics and supracellular organisation of cell shape fluctuations during dorsal closure. *Development*, 137(16):2743–2752, 2010.

## Abstract

Estimating the state of a dynamic system is an essential task for achieving important objectives such as process monitoring, identification, and control. Unlike linear systems, no systematic method exists for the design of observers for nonlinear systems. Although many researchers have devoted their attention to these issues for more than 30 years, there are still many open questions. We envisage that estimation plays a crucial role in biology because of the possibility of creating new avenues for biological studies and for the development of diagnostic, management, and treatment tools. To this end, this thesis aims to address two types of nonlinear estimation techniques, namely, the high-gain observer and the moving-horizon estimator with application to three different biological plants.
After recalling basic definitions of stability and observability of dynamical systems and giving a bird's-eye survey of the available state estimation techniques, we are interested in the high-gain observers. These observers may be used when the system dynamics can be expressed in specific a coordinate under the so-called observability canonical form with the possibility to assign the rate of convergence arbitrarily by acting on a single parameter called the high-gain parameter. Despite the evident benefits of this class of observers, their use in real applications is questionable due to some drawbacks: numerical problems, the peaking phenomenon, and high sensitivity to measurement noise. The first part of the thesis aims to enrich the theory of high-gain observers with novel techniques to overcome or attenuate these challenging performance issues that arise when implementing such observers. The validity and applicability of our proposed techniques have been shown firstly on a simple one-gene regulatory network, and secondly on an SI epidemic model.
The second part of the thesis studies the problem of state estimation using the moving horizon approach. The main advantage of MHE is that information about the system can be explicitly considered in the form of constraints and hence improve the estimates. In this work, we focus on estimation for nonlinear plants that can be rewritten in the form of quasi-linear parameter-varying systems with bounded unknown parameters. Moving-horizon estimators are proposed to estimate the state of such systems according to two different formulations, i.e., "optimistic" and "pessimistic". In the former case, we perform estimation by minimizing the least-squares moving-horizon cost with respect to both state variables and parameters simultaneously. In the latter, we minimize such a cost with respect to the state variables after picking up the maximum of the parameters. Under suitable assumptions, the stability of the estimation error given by the exponential boundedness is proved in both scenarios.
Finally, the validity of our obtained results has been demonstrated through three different examples from biological and biomedical fields, namely, an example of one gene regulatory network, a two-stage SI epidemic model, and Amnioserosa cell's mechanical behavior during Dorsal closure.



Keywords: nonlinear observers; high-gain observer, moving-horizon estimator (MHE); linear matrix inequalities (LMIs); Lyapunov stability; ISS stability; biological plants.



## Résumé

L'estimation de l'état d'un système dynamique est une étape cruciale, que ce soit pour la synthèse d'un contrôleur ou simplement pour l'identification ou la surveillance des processus. Une façon usuelle de résoudre ce problème consiste à implémenter un algorithme, appelé observateur, dont le rôle est de déduire une estimation fiable de l'état complet du système. Contrairement aux systèmes linéaires, aucune méthode systématique n'existe pour la conception d'observateurs pour les systèmes non linéaires. Bien que de nombreux chercheurs se soient penchés sur ces questions depuis plus de 30 ans, de nombreuses questions restent ouvertes. Nous envisageons que l'estimation joue un rôle crucial en biologie en raison de la possibilité de créer de nouvelles avenues pour les études biologiques et pour le développement d'outils de diagnostic, de gestion et de traitement. À cette fin, cette thèse vise à aborder deux types de

techniques d'estimation non linéaires, à savoir l'observateur à grand gain et l'estimateur à horizon mobile avec application à trois modèles biologiques.
Après avoir rappelé quelques concepts fondamentaux sur la stabilité et l'observabilité des systèmes dynamiques, puis passer en revue les principales techniques d'estimation d'état disponibles dans la littérature, nous nous intéressons aux observateurs à grand gain. Ces observateurs peuvent être utilisés lorsque la dynamique du système peut être exprimée en coordonnée spécifique sous la forme canonique dite d'observabilité avec la possibilité d'attribuer arbitrairement le taux de convergence en agissant sur un seul paramètre appelé paramètre de gain élevé. Malgré les avantages évidents de cette classe d'observateurs, leur utilisation dans des applications réelles est douteuse en raison de certains inconvénients : problèmes numériques, le problème de peaking et sensibilité élevée au bruit de mesure. La première partie de la thèse vise à enrichir la théorie des observateurs à grand gain avec de nouvelles techniques pour surmonter ou atténuer ces problèmes de performances difficiles qui surviennent lors de la mise en œuvre de tels observateurs.
La deuxième partie de la thèse étudie le problème de l'estimation d'état en utilisant l'approche d'estimation à horizon glissant (MHE). Le principal avantage du MHE est que les informations sur le système peuvent être explicitement considérées sous la forme de contraintes et donc améliorer les estimations. Dans ce travail, nous nous concentrons sur l'estimation des modèles non-linéaires qui peuvent être réécrits sous la forme de systèmes quasi-linéaires à paramètres variants dont des paramètres inconnus sont bornés. Des estimateurs à horizon glissant sont proposés pour estimer l'état de tels systèmes selon deux formulations différentes, à savoir "optimiste" et "pessimiste". Dans le premier cas, nous effectuons une estimation en optimisant le coût au sens des moindres carrés par rapport aux variables d'état et aux paramètres simultanément. Dans l'approche dite "pessimiste", on optimise un tel coût par rapport aux variables d'état après avoir pris le maximum des paramètres. Sous des hypothèses appropriées, la stabilité de l'erreur d'estimation donnée par la délimitation exponentielle est prouvée dans les deux scénarios.
Enfin, la validité de nos résultats obtenus a été démontrée à travers trois exemples différents issus des domaines biologiques et biomédicaux, à savoir un exemple d'un réseau de régulation génétique, un modèle épidémique de classe SI, et enfin le comportement mécanique des cellules Amniosereuse lors de la fermeture dorsale.

Mots- clés : observateurs non linéaires ; observateur grand gain; estimation à horizon glissant; inégalités matricielles linéaires (LMIs) ; stabilité de Lyapunov, stabilité ISS; modèles biologiques.

# Sommario

La stima di uno stato di un sistema dinamico è un passaggio essenziale per ottenere obiettivi importanti quali monitoraggio, identificazione e controllo del processo. Contrariamente ai sistemi lineari, non vi è un metodo di sistema per la valutazione per gli osservatori di uno non lineare. Sebbene molti ricercatori abbiano dedicato i loro studi a questi argomenti per più di 30 anni, ci sono tuttora molti punti da approfondire. Prevediamo che questo studio giochi un ruolo cruciale sulla biologia a causa della possibilità di creare nuove strade per gli studi sulla biologia e sullo sviluppo di diagnostica, management e attrezzatura di trattamento. Per concludere, questa tesi ha lo scopo di far emergere le tecniche di studio di due sistemi non lineari, ossia per l'osservatorio ad alto guadagno e lo studio con l'orizzonte in movimento con l'applicazione di tre differenti modelli biologici. Dopo aver riproposto le definizioni base di stabilità e osservabilità di sistemi dinamici, e aver ripreso le tecniche sugli studi disponibili, ci siamo focalizzati sugli osservatori ad alto guadagno. Questi osservatori possono essere impiegati quando i sistemi dinamici possono essere espressi in una specifica coordinata sotto la canonica forma dell'osservabilità, con la possibilità di assegnare la velocità di convergenza arbitrariamente agendo su un singolo parametro, chiamato, appunto, osservatorio ad alto guadagno. Nonostante gli evidenti vantaggi di questa classe di osservatori, il loro uso sulle applicazioni reali è questionabile a causa di alcuni inconvenienti: problemi numerici, il fenomeno di picco e l'alta sensibilità a misurare il rumore. La prima parte della tesi ha lo scopo di arricchire la teoria degli osservatori ad alto guadagno con nuove tecniche per superare o mitigare questi cambiamenti di rendimento che si raggiungono quando si implementano questi osservatori. La validità e l'applicazione delle nostre tecniche proposte sono state mostrate dapprima in un semplice lavoro di rete su un singolo gene e successivamente su un modello epidemico SI. Il principale vantaggio di MHE è che l'informazione sul sistema può essere esplicitamente considerato nella forma di vincoli e quindi migliorare le stime. In questo lavoro ci focalizziamo sulla stima di modelli non lineari che possono essere riscritti sotto forma di sistemi variabili parametrici quasi lineari con incognita limitata. I criteri dell'orizzonte mobile sono proposti per valutare lo stato di tali sistemi in base a due diverse simulazioni, come ad esempio "ottimistico" o "pessimistico". Nel primo caso eseguiamo la stima minimizzando il costo dell'orizzonte mobile dei minimi quadrati contemporaneamente sia rispetto alle variabili di stato che ai parametri. Nel secondo, invece, riduciamo al minimo tale costo rispetto alle variabili di stato dopo avere ottimizzato i parametri stessi. Sotto opportune ipotesi, la stabilità dell'errore di stima, data dalla limitatezza esponenziale, è dimostrata in entrambi gli scenari. Per concludere, la validità dei nostri risultati ottenuti è stata dimostrata attraverso tre diversi esempi, dal campo biologico e biomedico, ossia un esempio di una rete di regolazione genica, un modello epidemico SI a due stadi e il comportamento meccanico delle cellule di Amnioserosa durante la chiusura dorsale.

Parole chiave : osservatori; osservatori ad alto guadagno; stima dell'orizzonte mobile; disuguaglianze di matrici lineari (LMIs); stabilità di Lyapunov, stabilità ISS; modelli biologici.