On quantization and sporadic measurements in control systems: stability, stabilization, and observer design

Francesco Ferrante

THÈSE

En vue de l'obtention du

DOCTORAT DE L'UNIVERSITÉ DE TOULOUSE

Délivré par :
l'Institut Supérieur de l'Aéronautique et de l'Espace (ISAE)

Présentée et soutenue par:
Francesco Ferrante
le *21/10/2015*

On quantization and sporadic measurements in control systems: stability, stabilization, and observer design

JURY

Jean-Marc Biannic	Directeur de recherche	Examinateur
Paolo Frasca	Assistant professor	Examinateur
Antoine Girard	Directeur de recherche	Rapporteur
Frédéric Gouaisbaut	Maître de conférences	Examinateur
Samson Lasaulce	Directeur de recherche	Invité
Daniel Liberzon	Full professor	Rapporteur
Christophe Prieur	Directeur de recherche	Examinateur
Ricardo G. Sanfelice	Associate professor	Examinateur
Sophie Tarbouriech	Directeur de recherche	Examinateur

École doctorale et spécialité :
EDSYS : Automatique 4200046
Unité de Recherche :
Laboratoire d'analyse et d'architecture des systèmes (UPR 8001)
Directeur(s) de Thèse :
Sophie Tarbouriech et *Frédéric Gouaisbaut*
Rapporteurs :
Daniel Liberzon et *Antoine Girard*

This page is intentionally left blank.

Université de Toulouse

Institut Supérieur de l'Aéronautique et de l'Espace

On Quantization and Sporadic Measurements in Control Systems: Stability, Stabilization, and Observer Design

A Dissertation submitted in partial satisfaction of the requirements for the degree of

Doctor of Philosophy

in

Control Engineering

by

Francesco Ferrante

October 2015

This page is intentionally left blank.

On Quantization and Sporadic Measurements in Control Systems: Stability, Stabilization, and Observer Design

by

Francesco Ferrante

This manuscript has been composed in LaTeX

This page is intentionally left blank.

To my wife, Viola
To my parents, Aldo and Luana
To my sister, Marta

This page is intentionally left blank.

La filosofia è scritta in questo grandissimo libro che continuamente ci sta aperto innanzi a gli occhi (io dico l'universo), ma non si può intendere se prima non s'impara a intender la lingua, e conoscer i caratteri, nè quali è scritto. Egli è scritto in lingua matematica, e i caratteri son triangoli, cerchi, ed altre figure geometriche, senza i quali mezzi è impossibile a intenderne umanamente parola; senza questi è un aggirarsi vanamente per un oscuro laberinto.

Galileo Galilei *G. Galilei, Il Saggiatore, VI, 232*

This page is intentionally left blank.

ACKNOWLEDGMENTS

I would like to thank my advisors and friends Sophie Tarbouriech and Frédéric Gouaisbaut, who led me with patience, wisdom, and comprehension towards the completion of my doctorate program, while giving me intellectual freedom I needed. Their guidance has been a valuable support all along my doctoral program and throughout the writing of this thesis. I really learned a lot from them and I really liked the time spent discussing at the whiteboard transforming ideas in papers, I am sure I will miss all of that.

I want to express my deep gratitude to Ricardo G. Sanfelice, he has always encouraged me to pursue high quality research and taught me to never give up. I really benefited from his valuable advice and feedback on my research. I am indebted to him also for welcoming me at the University of California Santa Cruz during my visiting period and to be willing to serve as a committee member for my doctoral defense.

I am extremely thankful to Daniel Liberzon and Antoine Girard for having reviewed my thesis. It has been an honor for me.

I would like to sincerely thank Jean-Marc Biannic, Paolo Frasca, Antoine Girard, Samson Lasaulce, and Christophe Prieur for having accepted to serve as members for my doctoral committee.

My gratitude also goes to my mentor and friend Luca Zaccarian. He has always been an example and a source of inspiration, as well as a great teacher.

Special thanks go to my friend and office mate Simone Naldi. I will really miss the time spent discussing about math and enjoying the doctoral life.

I would like to thank also my friends Alessandro, Fuccio, Sebastiano, Amelia, Said, Luca and Valentina who made me feel literally like at home.

My sincere thanks go to the members of the MAC group who welcomed me in the best way and supported my attendance to various conferences.

I am thankful to all the faculty of the EEA department of the Université Paul Sabatier who assisted me during my teaching activities.

I am deeply grateful to the French Government for providing the funding sources that allowed me to pursue my doctorate program.

I want to express my deep gratitude also to Susy Baldini and her family, without you this never would have been possible.

Special thanks go to Sara Roe and her family who welcomed me at Santa Cruz as a family making my stay in the US a unique experience.

I would like to deeply thank my family that always encouraged me to pursue my objectives with commitment.

My last and foremost thanks go to my wife Viola, who constantly provided me with the love and the confidence needed to succeed in this endeavor, thanks my love.

Abstract

On Quantization and Sporadic Measurements in Control Systems: Stability, Stabilization, and Observer Design

by

Francesco Ferrante

In this dissertation, two fundamental aspects arising in modern engineered control systems will be addressed: On the one hand, the presence of quantization in standard control loops. On the other hand, the state estimation in the presence of sporadic available measurements. These two aspects are addressed in two different parts. One of the main feature of this thesis consists of striving to derive computer-aided tools for the solution to the considered problems. Specifically, to meet this requirement, we revolve on a linear matrix inequalities (LMIs) approach.

In the first part, we propose a set of LMI-based constructive Lyapunov-based tools for the analysis and the design of quantized control systems involving linear plants and linear controllers. The entire treatment revolves on the use of differential inclusions as modeling tools and on stabilization of compact sets as a stability notion.

In the second part of the thesis, inspired by some of the classical observation schemes presented in the literature of sampled-data observers, we propose two observers to exponentially estimate the state of a linear system in the presence of sporadic measurements. In addition, building upon one of the two observers, an observer-based controller architecture is proposed to asymptotically stabilize a linear plant in the presence of sporadic sensing and actuation.

This page is intentionally left blank.

Résumé

Sur la quantification et l'intermittence de mesures dans les systèmes de commande: stabilité, stabilisation, et estimation d'état.

par

Francesco Ferrante

Dans cette thèse, nous aborderons deux aspects fondamentaux qui se posent dans les systèmes de commande modernes du fait de l'interaction entre des processus en temps continu et des dispositifs numériques: la synthèse de lois de commande en présence de quantificateurs et l'estimation d'état en présence de mesures sporadiques. Une des caractéristiques principales de cette thèse consiste également à proposer des méthodes constructives pour résoudre les problèmes envisagés. Plus précisément, pour répondre à cette exigence, nous allons nous tourner vers une approche basée sur les inégalités matricielles linéaires (LMI).

Dans la première partie de la thèse, nous proposons un ensemble d'outils constructifs basés sur une approche LMI, pour l'analyse et la conception de systèmes de commande quantifiés impliquant des modèles et des correcteurs linéaires. L'approche est basée sur l'utilisation des inclusions différentielles qui permet de modéliser finement le comportement de la boucle fermée et ainsi d'obtenir des résultats intéressants.

Dans la seconde partie de la thèse, inspirés par certains schémas d'observation classiques présentés dans la littérature, nous proposons deux observateurs pour l'estimation de l'état d'un système linéaire en présence de mesures sporadiques, c'est-à-dire prenant en compte la nature discrète des mesures disponibles. De plus, en se basant sur une des deux solutions présentées, une architecture de commande basée observateur est proposée afin de stabiliser asymptotiquement un système linéaire en présence à la fois de mesures sporadiques et d'un accès intermittent à l'entrée de commande du système.

This page is intentionally left blank.

CONTENTS

GENERAL INTRODUCTION

In this dissertation, two fundamental aspects arising in modern engineered control systems will be addressed: On the one hand, the presence of quantization in standard control loops. On the other hand, the state estimation in the presence of sporadic available measurements. These two aspects are addressed in two different parts.

One of the main feature of this thesis consists of striving to derive computer-aided tools for the solution to the considered problems. Specifically, to meet this requirement, we revolve on a linear matrix inequalities (LMIs) approach. The spirit of such an approach consists of formulating the considered problem directly in a form that is convenient from a numerical standpoint, instead to derive closed form solutions, which can be a cumbersome, often impossible, challenge. Then, thanks to the availability of efficient algorithms for the solutions of LMIs, the solution to the considered problem can be derived through efficient computer-aided tools; see, *e.g.*, [126] for an interesting survey on this aspect.

The contents of the two parts composing this thesis are briefly illustrated below.

Quantization in control system

Most of the modern engineered systems are composed by continuous-time plants interacting with digital devices and/or data networks. In all these settings, quantization is an always present phenomenon, *e.g.*, [17, 21, 32, 35, 51, 84, 116, 117] just to cite a few.

In this first part of this thesis, we propose a set of LMI-based constructive Lyapunov-based tools for the analysis and the design of quantized control systems involving linear plants and linear controllers. The entire treatment revolves on the use of differential inclusions as modeling tools, and on stabilization of compact sets as a stability notion.

State estimation and observer-based control in the presence of sporadic measurements

In real-world engineering applications, assuming to continuously measuring the output of a given plant is undoubtedly unrealistic. This practical needed has brought to life a new research area aimed at developing observer schemes accounting the discrete nature of the available measurements; see, *e.g*, [1, 4, 6, 74, 92].

In this part of this thesis, inspired by some of the classical observation schemes presented in the literature of sampled-data observers, we propose two observers to exponentially estimate the state of a linear system in the presence of sporadic measurements. In addition, building upon one of the two observers, an observer-based controller architecture is proposed to asymptotically stabilize a linear plant in the presence of sporadic measurements and intermittent input access. The design of such a controller is streamlined by the derivation of a separation principle for the considered architecture.

A unique feature of the proposed approach consists of hinging upon the hybrid systems framework proposed in [56]. On the one hand, by following this approach a completely novel modeling of the considered observers is provided, as well as the derivation of novel systematic design strategies is illustrated. On the other hand, the huge flexibility provided by the framework in [56] allows to envision very appealing extensions of the results presented in this part, giving rise to novel lines of research.

Part I

Quantization in Control Systems

INTRODUCTION

General Overview and some Historical Aspects

Recently technology enhancements have enabled the conception of a new generation of engineered systems integrating physical interactions, computational and communication abilities. The rapid spreading of this kind of systems stems from the worthy advantages in scalability, ease of maintenance and high computational resources entailed by the use of cutting-edge technology solutions in real-world applications, as transportation systems, automotive, autonomous robotics, energy delivery systems etc. This new trend has been having a strong impact also in modern control systems that are nowadays built via the adoption of digital controllers and digital instrumentation [93]. Typically physical systems evolve continuously as the ordinary time flows and are characterized by variables that take values in uncountable sets. Instead, digital devices evolve in a discrete fashion and their evolution is characterized by variables taking values in countable set. When a physical system interacts with a digital one, side effects as time-delays, asynchronism, quantization, are unavoidable issues that can often turn into an overblown performance degradation, like the appearing of limit cycles or chaotic phenomena or even instability of the closed-loop system.

Concerning the effect of quantization in control systems, since such a phenomenon is almost pervasive in modern engineered control systems, its study has extensively attracted researchers over the last years; see, *e.g.*, [17, 21, 32, 35, 51, 84, 116, 117] just to cite a few.

The negative impact of quantization on control systems seems to be already known in the late 50's, an attempt to tackle with this phenomenon can be traced back in the work of Kalman featured in [70]. In this paper, quantization was essentially addressed via stochastic tools. In fact, until the late 80's, the common trend considered by researchers in addressing quantization in control systems consisted to look at quantization as a phenomenon inducing a non deterministic deviation of the quantized control system from its nominal (quantization free) behavior. Therefore, the standard custom was designing controllers via standard techniques while overlooking quantization. Then, to somehow to capture the real behavior of the closed-loop suitable stochastic characterizations of the quantization error were considered; see [6]. Clearly this approach can be effective whenever the level of specification is rather modest and the quantization somehow restrained. Therefore, since digital devices at that time were becoming pervasive in control systems and at the same time the level of performances required was continuously increasing, new systematical tools to deal with quantized control systems in their actual nature were necessary. In the late 80's, the works of Delchamps [34, 35], and to some extent the one of Miller et al. [89], marked a watershed in the literature of quantized control systems proposing an alternative approach to deal with stability and stabilization in quantized control systems. Such an approach consists of modeling the quantization phenomenon through a static nonlinear function, the *quantizer*, mapping a real variable into a variable belonging to a countable set $\mathcal{Q}$, *i.e.*, $\mathrm{q}\colon \mathbb{R} \to \mathcal{Q}$. The methodology proposed by Delchamps et al. ([34, 35]) is relevant since it has brought to life a new research area founded on the tools issued from the nonlinear control theory for the study of quantized control systems. From then, the rapid development of the control systems science in the setting of quantized control has rapidly given rise to different approaches and tools to deal with quantization in control systems. Essentially such approaches share a common fundamental idea that builds on a robust control point of view. Namely, the closed-loop system is modeled as a nominal system perturbed by a (potentially locally) bounded perturbation, *i.e.*, the quantization error. First attempts resting on this approach for the special case of SISO systems can be found in [89]. In particular, in [89] the authors attack the problem of having quantized measurements in a linear control system by first bounding the quantization error and then by pursuing a Lyapunov approach to establish ultimate boundedness. One of the main important feature of this paper consists of pointing out that asymptotic stability of the origin of quantized control systems can be unlikely achieved due to finite precision information provided by quantizers. Later on, this general approach has been extended in [17] to general linear systems with quantized measurements, in [82] to nonlinear systems in the presence of quantized control inputs or quantized measurements, while in [83] an observer-based controller architecture is presented to build an output feedback controller in the presence of quantized measurements. The key idea adopted by the authors in all these latter publications consists of addressing quantized control system via the input-to-state stability notion due to Sontag; see, *e.g.*, [114]. In particular, the authors shown that input-to-state stable control systems have the needed robustness to tolerate quantization. We emphasize that in all these works, the authors besides pointing out the relevance of input-to-state stability in quantized control systems, by relying on a more sophisticated type of quantizer allowing the possibility to dynamically scaling the quantization error (called in general dynamic quantizer), provided novel control policies to ensure asymptotic stabilization rather than ultimate boundedness. This approach has given rise to a complete novel line of research more focused on an information point of view, that is aimed at characterizing the quantity of information actually needed to achieve stabilization of a given plant depending on its open-loop behavior; see, *e.g.*, [121] and the references therein.

Subsequently, in [21] the authors by restricting the attention to logarithmic quantization and by pursuing a sector bound approach relax the input-to-state stability requirement to achieve stabilization of nonlinear systems with input quantization, at least for the case of logarithmic quantization. This fact of encapsulating quantization error into a sector before being used in [21] was already considered in [51], for the case of discrete-time linear systems. These latter approaches show that quantization can be effectively faced by the use of robust control tools as the sector bound approach. The main effort made in these latter works is concerned to achieve asymptotic stability of the origin via a quantizer as coarse as possible. On the other hand, the asymptotic stabilization of the origin can be achieved in general only when the considered quantizer is infinitely precise close to the origin, as it is for the case of the logarithmic quantizer. However, in some real-world settings the availability of such a kind of quantizers cannot be considered due to technological or optimization constraints. This consideration originated a complete analysis in [21, 36] of the case of finite symbols logarithmic quantizers. Specifically, in [21] the authors shown that in such a case under analogous conditions as in the case of the genuine logarithmic quantizer, semi-global practical stabilization can be easily achieved in the presence of a finite number of symbols, at least for the case on input quantization.

Another interesting and fundamental aspect linked to quantized control systems regards the issues related to discontinuous behaviors induced by quantizers in standard control loops. Indeed, the fact that quantizers map uncountable sets into countable ones implies that quantizers are essentially discontinuous mappings. This fact has a serious impact when quantizers interact with dynamical systems. Indeed, it is well known that discontinuities give rise to serious problems when coupled with differential or difference equations [22, 31, 46, 75, 78]. Such problems range from questions related to the existence and the nature of the solutions to the resulting closed-loop system (in continuous-time dynamical systems) to robustness issues of the closed-loop system with respect to small perturbation and/or measurement noise (continuous-time and discrete-time dynamical systems). The serious questions arising from discontinuities in differential equations were already known in the late 60's by the community working on differential equations, as testified by the work of Hájek in 1979 [59] that offers an interesting survey on this appealing topic.

Later on, the increasing number of real applications concerning discontinuous differential equations has notably boosted the research in this area. Such an intense research has led to a comprehensive and solid theory to address discontinuous right-hand side differential equations, important results and contributions in this field can be found in [10, 46, 77] just to cite a few, while an interesting and with a modern flavor survey on discontinuous dynamical systems is contained in [31]. We emphasize that the huge development of the modern theory of discontinuous dynamical systems have been made possible by the development of the theory of differential inclusions; see, *e.g.*, [7, 28], which are the main tool, although not the unique, to address discontinuous dynamical systems.

Despite the deep knowledge available nowadays about discontinuous dynamical systems, surprisingly no much work in that setting has been done in the literature to deal with quantized control systems. Building on the tools originally proposed in [77], a first work offering a treatment of quantized continuous-time control systems seems to appear in [21]. Further results have been presented later in [22].

In our opinion, the main reason behind this lack of contributions looking at quantized systems as discontinuous dynamical systems is mainly due to the fact that the greatest number of publications within this field deal with discrete-time systems rather than continuous-time ones. In the case of discrete-time systems, certainly the concerns related to the existence of solution are no longer a problem. Nonetheless, discontinuities in discrete-time systems may jeopardize the robustness of the resulting closed-loop system. Interesting examples about this aspect are shown, *e.g.*, in [75, 78]. On the one hand, pursuing a robust control approach, as the sector bound approach discussed above, generally prevents from running into poorly robust control systems even if the discontinuity is not directly accounted. On the other hand, such a discontinuity may give rise to behaviors for which a traditional analysis cannot provide any precise justification.

Nevertheless, in modern engineered systems the classical paradigm of considering quantization only paired with discrete-time systems needs to be reconsidered. Many examples can be found in which continuous-time dynamical systems interact with quantized variables; see, *e.g.*, [22]. Thus, a proper treatment of the situations falling into this context is a real need.

Contribution

The contribution we offer in this first part of this dissertation aims at bridging the gap left by the existing literature concerning the (almost) lack of constructive methods for quantized linear control systems, with a special focus on uniform quantization. Specifically, we restrict our interest to the class of continuous-time linear time-invariant systems. The issues related in having closed-loop systems modeled via discontinuous right-hand side differential equations will be faced via the proper tools proposed by literature, likewise to [21]. In particular, inspired by the literature of saturating systems, we provide constructive LMI-based conditions for the stability analysis and the controller synthesis encompassing several settings naturally arising in real-world applications. Such conditions enable to couple optimization aspects with the considered problems, in a similar, although dual, fashion to the case of saturated closed-loop systems. The use of optimization as a tool for conservatism reduction and closed-loop behavior improvement are the main aims of this thesis.

The main feature of the methodology we propose in this dissertation consists of merging together aspects arising from discontinuous-right hand side differential equations with a constructive approach.

The remainder of this part is organized as follows.

- Chapter 1 illustrates the modeling framework adopted in this dissertation to deal with quantized control systems, with a special emphasis on linear control systems and uniform quantization. Moreover, the technical foundations underlying the pursued approach are thoroughly illustrated and commented within this chapter.

- Chapter 2 deals with static state feedback control for linear systems in the presence of uniform quantization. In this setting, constructive conditions for the stability analysis and the controller design are provided. Some of the results presented in this chapter can be found in [40].
- Chapter 3 deals with dynamic output feedback control of linear systems in the presence of uniform quantization. Even in this case, the proposed approach is constructive and strives for obtaining tractable conditions from a numerical standpoint. Some of the results presented in this chapter are included in [37, 38].

Numerical solutions to LMIs throughout this dissertation are obtained via YALMIP [87] and coded in Matlab.

This page is intentionally left blank.

1

QUANTIZED CONTROL SYSTEMS: MODELING AND TECHNICAL FOUNDATIONS

"What is now proved was once only imagined."

– William Blake

1.1 Introduction

In this chapter, we present the quantization phenomenon in its general form, and the problems arising from the presence of quantizers in standard control loops. Then, the general aspects of quantization in control systems are sharpened for the case of linear control systems subject to uniform quantization. In this context, we illustrate some technical results, that will be used in the sequel of this dissertation.

1.2 Quantized Systems: Modeling

Following the general approach proposed in [35], in this dissertation, as *quantizer*, we mean a function q that maps the Euclidean space $\mathbb{R}^\ell$ into a countable set $\mathcal{Q} \subset \mathbb{R}^\ell$, that is:

$$\mathrm{q}\colon \begin{cases} \mathbb{R}^\ell \to \mathcal{Q} \\ x \mapsto \mathrm{q}(x). \end{cases} \tag{1.1}$$

In this part of this dissertation, we are interested in analyzing the impact of quantization on standard control systems. Specifically, let us consider the following nonlinear plant

$$\begin{cases} \dot{x} = f(x, u) \\ y = h(x) \end{cases} \tag{1.2}$$

where $x \in \mathbb{R}^n$ is the state, $u \in \mathbb{R}^m$ is the control input and $y \in \mathbb{R}^p$ is the plant output, that in some cases can also coincide with the whole state vector x. $f\colon \mathbb{R}^n \to \mathbb{R}^n$, and $h\colon \mathbb{R}^n \to \mathbb{R}^p$ are two given functions.

Suppose that the system (1.2) is controlled through a feedback controller, whose input coincides with the measure of the plant output y, and generates a control signal u_c which feeds (1.2). On the other hand, in real implementations, the plant and the controller are not directly connected together. Indeed, measurements of the plant output are gathered via physical sensors. In modern applications, often such sensors have a finite precision, *e.g.*, optical encoders, digital sensors, etc. In all these situations, the measured plant output sent to the controller is represented by means of a discrete set of values *i.e.*, is quantized. In the sequel, we will denote this case as *sensor quantization*. Fully analogous considerations hold for the input channel. In particular, the adoption of finite-resolution actuators, (as, *e.g.*, stepper motors), or finite precision realization of the controller entails a quantization of the control signal. In the sequel, we will denote this case as *actuator quantization*. Moreover, actuator and sensor quantization may also occur simultaneously. For instance, this situation occur in distributed control systems, where the physical interconnection between the controller and the plant is ensured by a finite-bandwidth communication channel; see Figure 1.1. Indeed, in such a situation, the communication channel prevents from sending infinite precision data from one end to the other; see [22, 62]. Thus, in these contexts, building from (1.2), the open-loop plant model to be considered for the analysis, but even for the design, of the control system should be as follows

$$\begin{cases} \dot{x} = f(x, u) \\ u = \mathrm{q}_u(u_c) \\ y_m = \mathrm{q}_y(h(x)) \end{cases} \tag{1.3}$$

where y_m, and u_c are, respectively, the measured output and the signal sent to the plant.

Remark 1.1. In the proposed model (1.3), the dynamics of sensors and actuators do not directly appear. On the other hand, such dynamics can either be neglected, whenever they are much more faster of those of the plant, or be incorporated either in the plant model, or in the controller model. Thus, the modeling framework given in (1.3) is without loss of generality.

Concerning the controller structure, depending on the availability of plant state, we consider two classes of controllers.

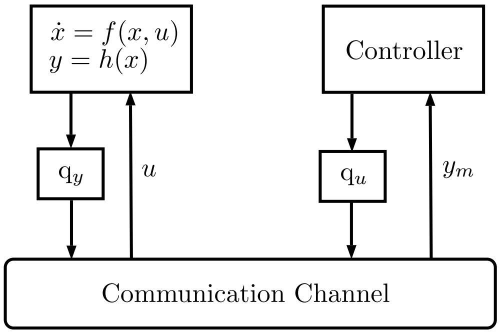


Figure 1.1: A networked control system. Both the controller and the plant communicate with the channel via a finite data rate.

Static State Feedback Controller

Whenever the plant state x is fully accessible, that is $h = \mathrm{id}$, we will adopt a static state feedback control law. In particular, in this setting, three situations can occur. In the first case, the plant state is assumed to be measured directly, that is $y_m = y = x$, and only the control input is subject to quantization. In this case, $u = \mathrm{q}_u(\kappa(x))$, where $\kappa \colon \mathbb{R}^n \to \mathbb{R}^m$ is a given function. In the second case, we assume that only the measured state x is quantized, which yields $u = u_c = \kappa(\mathrm{q}_y(x))$. Finally, in the third case, we assume that both the measured state and the control input are quantized, that is $u = \mathrm{q}_u(\kappa(\mathrm{q}_y(x)))$. In this latter case, that encompasses the two others, the closed-loop system reads

$$\begin{cases} \dot{x} = f(x, u) \\ u = \mathrm{q}_u(u_c) \\ u_c = \kappa(\mathrm{q}_y(x)) \end{cases} \tag{1.4}$$

Dynamic Output Feedback Controller

Whenever, the plant state is not fully accessible, we adopt a dynamic output feedback control law defined as follows

$$\begin{cases} \dot{x}_c = \eta(x_c, y_m) \\ u_c = \omega(x_c, y_m) \end{cases} \tag{1.5}$$

where $x_c \in \mathbb{R}^{n_c}$ is the controller state, and $\eta \colon \mathbb{R}^{n_c} \times \mathbb{R}^p \to \mathbb{R}^{n_c}$, $\omega \colon \mathbb{R}^{n_c} \times \mathbb{R}^p \to \mathbb{R}^m$ are two given functions. In this case, three different scenarios can be considered. In the first one, the plant output y is quantized, namely the measured output effectively accessible is $y_m = \mathrm{q}_y(y)$. In the second one, the control input u is quantized, namely $u = \mathrm{q}_u(u_c)$, while in the third one both the plant output, and the control input are quantized. In this latter case, that encompasses the two others, the closed-loop system reads

$$\begin{cases} \dot{x} = f(x,u) \\ \dot{x}_c = \eta(x_c, y_m) \\ u_c = \omega(x_c, y_m) \\ u = \mathrm{q}_u(u_c) \\ y_m = \mathrm{q}_y(h(x)) \end{cases} \tag{1.6}$$

1.2.1 Discontinuous Dynamical Systems

From the general representation given by (1.1), it turns out that a quantizer is a function that maps the Euclidean space into a countable set. This fact implies that, whatever is the way adopted to realize such a mapping, the resulting map is discontinuous. Recall that any continuous function maps the Euclidean space, which is connected, into a connected set, (see, *e.g.*, [107]), then in general uncountable[1]. Therefore, in any situation of those presented above, the closed-loop system is described by a discontinuous-right hand side differential equation. Therefore, there are no guarantees about the existence of classical solutions to the closed-loop system, *i.e.*, everywhere differentiable functions which satisfy the dynamics of the closed-loop system at each point in their domain; see [46]. To overcome this drawback, more general notions of solution are proposed in the literature. In particular, in this dissertation we will consider the notion of solution due to Carathéodory; see, *e.g.*, [22, 31], and the notion due to Krasovskii; [77]. In the sequel, such notions are thoroughly presented and illustrated in some examples. In particular, we introduce them for a dynamical system in the following form.

$$\dot{x} = X(x) \tag{1.7}$$

where $x \in \mathbb{R}^n$, and $X \colon \mathbb{R}^n \to \mathbb{R}^n$.

Definition 1.1 (Carathéodory solution, [31])**.** Let $\mathbb{I} \subset \mathbb{R}_{\geq 0}$ be an interval containing 0. A function $\varphi \colon \mathbb{I} \to \mathbb{R}^n$ is a Carathéodory solution to (1.7) if φ is absolutely continuous on $\mathbb{I}$, and[2]

$$\dot{\varphi}(t) = X(\varphi(t)) \qquad \text{for almost all} \quad t \in \mathbb{I}.$$

The above definition does not insist either on on the differentiability of φ or on the fact that (1.7) needs to be satisfied on the whole domain of the solution. This weakening with respect to the classical notion given by Peano ([98]) allows to deal with a wider class of situations often occurring in control problems.

To delve into this issue, let us consider the following example.

[1] The only countable connected sets are the singletons. But this case is not of interest in our setting

[2] Let $J \subset \mathbb{R}$ be a given interval, and $f \colon J \to \mathbb{R}^n$ be a given function, the derivatives of f are considered one-sided derivatives at the end points of J.

Example 1.1. Consider the system (1.7) for which

$$X(x) = \begin{cases} 1 & x = 0 \\ -1 & \text{elsewhere.} \end{cases} \tag{1.8}$$

Obviously, the system defined by (1.7)-(1.8) does not admit any solution φ in the sense given by Peano with $\varphi(0) = 0$, *i.e.*, a derivable function satisfying (1.7) for each $t \in \operatorname{dom} \varphi$. Indeed, let us assume that there exists such a solution φ defined over $[0, T]$, for some $T \in \mathbb{R}_{\geq 0}$. Then, since it needs to satisfy $\dot{\varphi}(0) = 1$, and being φ derivable, there would exist a small enough positive T', such that for every $t \in [0, T']$, $\dot{\varphi}(t) > 0$, giving $\varphi(t) > 0$ for $t \in (0, T']$. However, this contradicts the fact that φ satisfies (1.7)-(1.8) over $[0, T]$.

In the above example, the issue preventing from the existence of a classical solution φ, with $\varphi(0) = 0$, stems from the fact that the discontinuity of the right-hand side imposes a constraint that does not allow φ to flow away from zero. Obviously, this drawback only occurs whenever a solution comes across to the origin. In particular, completely different conclusions can be drawn by following the notion of solution due to Carathéodory. This fact is shown in the following example.

Example 1.2. Let us consider the system defined by (1.7)-(1.8). We want to investigate the existence of Carathéodory solutions, φ, with $\varphi(0) = 0$, to such a system. According to Definition 1.1, for every $T > 0$, $\varphi(t) = -t$ is a Carathéodory solution for (1.7)-(1.8), Indeed, such a solution is such that $\dot{\varphi}(t) = X(\varphi(t))$, for every $t \in (0, T]$. Namely, φ does not satisfy the related differential equation in $t = 0$, *i.e.*, it satisfies (1.7)-(1.8) for almost all $t \in [0, T]$.

The above two examples have the merit to show how via a more general notion of solution, one may overcome drawbacks arising from discontinuous right-hand side differential equations. However, in some cases, the notion of solution due to Carathéodory is not weak enough to guarantee the existence of solutions. To understand the relevance of this issue, let us consider the following example, which situates more in the context of this dissertation.

Example 1.3. Consider the following given plant with quantized actuator

$$\begin{cases} \dot{x} = u \\ u = \mathrm{q}(u_c) \end{cases}$$

Specifically, $\mathrm{q}\colon u_c \mapsto \operatorname{sign}(u_c)$, for which we consider $\operatorname{sign}(0) = 1$. That is q maps the Euclidean space into $\{-1, 1\}$.

Let us suppose that we want to stabilize the above plant via the following static state feedback controller $u_c = -x$. Then, the closed-loop system reads

$$\dot{x} = -\operatorname{sign}(x). \tag{1.9}$$

Clearly, the closed-loop system does not admit any Carathéodory solution φ with $\varphi(0) = 0$. Indeed, by contradiction, let φ be a Carathéodory solution to (1.9) with $\varphi(0) = 0$. For every $x \in \mathbb{R}$, define the function $W(x) = \frac{1}{2}x^2$. Then, since φ is absolutely continuous on its domain, and $W(x)$ is continuously differentiable on $\mathbb{R}$, the function $W(\varphi(t))$ is absolutely continuous on $\operatorname{dom}\varphi$. Hence, its derivative exists for almost all $t \in \operatorname{dom}\varphi$, and whenever it exists

$$\frac{dW(\varphi(t))}{dt} = -\varphi(t)\,\mathrm{sign}(\varphi(t)) = -|\varphi(t)|.$$

Thus, since $W(\varphi(t))$ is absolutely continuous, then

$$W(\varphi(t)) = \int_0^t \frac{dW(\varphi(s))}{ds}ds = -\int_0^t |\varphi(s)|ds \qquad \forall t \in \operatorname{dom}\varphi$$

where the above integral needs to be intended as a Lebesgue integral; see, *e.g.*, [107].

Now, as $W(x)$ is nonnegative for every $x \in \mathbb{R}$, then for almost all $t \in \operatorname{dom}\varphi$, it needs to be $\varphi(t) = 0$. But, such a function is not a Carathéodory solution to (1.9). Indeed, suppose that φ is a solution to (1.9), and that it is equal to zero for almost all $t \in \operatorname{dom}\varphi$. Then, it follows that,

$$\varphi(t) = -\int_0^t \mathrm{sign}(\varphi(s))ds = -t \qquad \forall t \in \operatorname{dom}\varphi$$

but this contradicts the fact that φ is equal to zero for almost all $t \in \operatorname{dom}\varphi$.

The above example shows that unfortunately the notion of solution due to Carathéodory is still not enough to guarantee the existence of solutions for a given discontinuous right-hand side differential equation. To overcome this problem, in the literature several notions of solution are proposed; see, *e.g.*, [8, 46, 77, 111]. In this dissertation, we embrace the notion of solution due to Krasovskii [77].

Definition 1.2 (Krasovskii solution [59])**.** For each $x \in \mathbb{R}^n$, let us define the following set-valued mapping

$$\mathcal{K}[X](x) := \bigcap_{\delta>0} \overline{\mathrm{co}}X(x + \delta\mathbb{B}) \tag{1.10}$$

where $\mathbb{B}$ is the closed unitary ball in $\mathbb{R}^n$. A function $\varphi\colon \mathbb{I} \to \mathbb{R}^n$, with $\mathbb{I} \subset \mathbb{R}_{\geq 0}$ an interval containing 0, is a Krasovskii solution to (1.7) if it is absolutely continuous on $\mathbb{I}$, and

$$\dot{\varphi}(t) \in \mathcal{K}[X](\varphi(t)) \qquad \text{for almost all} \quad t \in \mathbb{I}.$$

In this dissertation, for any function X, we will refer to the set-valued mapping $\mathcal{K}[X](x)$ as Krasovskii regularization of X (this terminology is proposed in [56]).

Remark 1.2. Notice that, the Krasovskii regularization of a locally bounded function $X\colon \mathbb{R}^n \to \mathbb{R}^n$ has some interesting properties as set-valued mapping. In particular, by definition of the Krasovskii regularization, it follows that for each $x \in \mathbb{R}^n$, $\mathcal{K}[X](x)$ is convex, $\operatorname{dom}\mathcal{K}[X] = \mathbb{R}^n$, and according to [56, Lemma 5.16] $\mathcal{K}[X]$ is outer semicontinuous. In addition, local boundedness of X yields local boundedness of $\mathcal{K}[X]$. These properties will be of interest in the sequel of this dissertation.

Three main reasons encourage to choice this kind of notion in control problems. The first one is that Krasovskii solutions exist under very mild requirements, (below a formal result concerning existence of Krasovskii solutions is given). The second one is that, whenever they exist, Carathéodory solutions are Krasovskii solutions. Then, any conclusion drawn on Krasovskii solutions also holds for Carathéodory solutions. The third one is that, as shown in [59, Corollary 5.6.], (and also more recently in [56, Theorem 4.3.]), Krasovskii solutions coincide with Hermes solutions, which are defined as follows

Definition 1.3 (Hermes solutions [59])**.** A function $\varphi_{\mathcal{H}}$ is a Hermes solution to (1.7) on a compact interval $J \subset \mathbb{R}_{>0}$, if there exist a sequence of measurable functions $\{p_k\}_{k=0}^{\infty}$ defined on J, and a sequence of functions $\{\varphi_k\}_{k=1}^{\infty}$ defined on J, such that φ_k is a Carathéodory solution to $\dot{\varphi}_k = X(\varphi_k + p_k)$, the sequence $\{p_k\}_{k=1}^{\infty}$ converges uniformly to the zero function on J, and φ_k converges uniformly to $\varphi_{\mathcal{H}}$ on J.

The notion of Hermes solutions allows to capture the effect of arbitrarily small state perturbations on the solutions to (1.7). Such perturbations may represent actuation disturbances, measurement noises, or modeling errors. Thus, this fact provides a strong justification fostering the adoption of Krasovskii (Hermes) solutions in control problems. The reader may consult [56, Example 4.1.] for a interesting example showing connections between Krasovskii solutions and Hermes solutions, in a case similar to Example 1.2. Concerning the existence of Krasovskii solutions, let us consider the following result given, *e.g.*, in [23, 56], and which is direct consequence of general results on differential inclusions presented in [7]. Such a result uses the notion of locally bounded function.

Definition 1.4 ([30])**.** A function $f\colon S$ is locally bounded if for every $s \in S$ there exists a neighborhood $\mathcal{B}$ of s, such that $f(\mathcal{B})$ is bounded.

Theorem 1.1. Let $x_0 \in \mathbb{R}^n$. If X is locally bounded, then there exists at least a Krasovskii solution φ to (1.7), such that $\varphi(0) = x_0$.

To exploit the notion of solution due to Krasovskii, one needs to compute the Krasovskii regularization of the function X, which in general is a nontrivial task. To simplify such a task, we illustrate below some properties of the Krasovskii regularization for a given function X. Such properties were originally proposed for the Filippov regularization in [97], and then extended to the Krasovskii regularization in [23].

Proposition 1.1.

(i) *If $X\colon \mathbb{R}^{\ell_1} \to \mathbb{R}^{\ell_2}$ is continuous at $x \in \mathbb{R}^{\ell_1}$, then $\mathcal{K}[X](x) = \{X(x)\}$*

(ii) *Given two locally bounded functions $X_1, X_2\colon \mathbb{R}^{\ell_1} \to \mathbb{R}^{\ell_2}$, then $\mathcal{K}[X_1{+}X_2](x) \subseteq \mathcal{K}[X_1](z) + \mathcal{K}[X_2](z)$. Moreover, if either X_1 or X_2 are continuous at $x \in \mathbb{R}^{\ell_1}$, then equality holds.*

(iii) *Given two locally bounded functions $X_1\colon \mathbb{R}^{\ell_1} \to \mathbb{R}^{\ell_2}$, and $X_2\colon \mathbb{R}^{\ell_1} \to \mathbb{R}^{\ell_3 \times \ell_2}$, ($X_2$ is a matrix valued function). If X_2 is continuous at $x \in \mathbb{R}^{\ell_1}$, then $\mathcal{K}[X_2X_1](x) = X_2(x)\mathcal{K}[X_1](x)$; where for every $x \in \mathbb{R}^{\ell_1}$, $X_2X_1(x) := X_2(x)X_1(x)$.*

Moreover, as follows, we propose another result, that will be of interest in the sequel. Such a result is somehow derived from [97].

Proposition 1.2. *Let $X_1\colon \mathbb{R}^{\ell_1} \to \mathbb{R}^{\ell_2}$ be a locally bounded function, and $X_2\colon \mathbb{R}^{\ell_3} \to \mathbb{R}^{\ell_1}$ a* *continuous function. Then, for each* $x \in \mathbb{R}^{\ell_3}$*,*

$$\mathcal{K}[X_1 \circ X_2](x) \subseteq \mathcal{K}[X_1](X_2(x)). \tag{1.11}$$

Proof. First of all, for each $x \in \mathbb{R}^{\ell_3}$, let us define the set[3]

$$\mathcal{L}(x) = \{\lim X_1 \circ X_2(x_k) | x_k \to x\} \subset \mathbb{R}^{\ell_2}$$

where x_k is any sequence converging to x. Since $X_1 \circ X_2$ is locally bounded, according to [9, Lemma 1], it turns out that for every $x \in \mathbb{R}^{\ell_3}$, $\mathcal{K}[X_1 \circ X_2](x) = \operatorname{co} \mathcal{L}(x)$. For each $x \in \mathbb{R}^{\ell_3}$, define the set

$$\mathcal{P}(x) = \{\lim X_1(p_k) | p_k \to X_2(x)\} \subset \mathbb{R}^{\ell_2}$$

where p_k is any sequence converging to $X_2(x)$. Pick any $l \in \mathcal{L}(x)$, by definition, there exists a sequence $x_k \to x$, such that $l = \lim X_1 \circ X_2(x_k)$. For any $k \in \mathbb{N}$, define the sequence $\tilde{p}_k = X_2(x_k)$, then $l = \lim X_1(\tilde{p}_k)$. On the other hand, since X_2 is continuous, then $\tilde{p}_k \to X_2(x)$, which implies that $l \in \mathcal{P}(x)$. Since this property holds for any $l \in \mathcal{L}(x)$, it follows that, for each $x \in \mathbb{R}^{\ell_3}$,

$$\mathcal{L}(x) \subseteq \mathcal{P}(x)$$

Therefore, taking the convex-hull of both sides of the above relation and recalling that for each $x \in \mathbb{R}^{\ell_3}$

$$\mathcal{K}[X_1](X_2(x)) = \operatorname{co} \mathcal{P}(x)$$

establishes the result. ■

Remark 1.3. Notice that showing the complementary inclusion to (1.11) requires additional assumptions on the function X_2. In particular, the equality can be established requiring that X_2 is smooth and that for each $x \in \mathbb{R}^{\ell_3}$ $\operatorname{rank} \nabla X_2(x) = \ell_1$; see [97].

Another result, still derived from [97], is given next. Such a result is useful to address decentralized discontinuous functions, often occurring in control problems.

Proposition 1.3. *For each* $i = 1, 2, \ldots, \ell$*, let* $X_i \colon \mathbb{R}^{n_i} \to \mathbb{R}^{n_i}$ *be locally bounded functions. Let, for each* $x \in \bigtimes_{i=1}^{\ell} \mathbb{R}^{n_i}$*,*

$$Y(x) = \bigtimes_{i=1}^{\ell} X_i(x_i).$$

Then, for each $x \in \bigtimes_{i=1}^{\ell} \mathbb{R}^{n_i}$*, the following identify holds*

$$\mathcal{K}[Y](x) = \bigtimes_{i=1}^{\ell} \mathcal{K}[X_i](x_i). \tag{1.12}$$

Proof. For notation simplicity, we prove the above result for $\ell = 2$, the extension to the

[3] This notation is inherited by the seminal work of Paden and Sastry [97] presenting calculation rules for the Filippov regularization.

general case is straightforward. First of all, for each $x \in \mathbb{R}^{n_1} \times \mathbb{R}^{n_2}$, let us define the set

$$\mathcal{L}(x) = \{\lim Y(x_k) | x_k \to x\} \subset \mathbb{R}^{n_1} \times \mathbb{R}^{n_2}$$

where x_k is any sequence converging to x. Since Y is locally bounded, according to [9, Lemma 1], it turns out that for every $x = (x_1, x_2) \in \mathbb{R}^{n_1} \times \mathbb{R}^{n_2}$,

$$\mathcal{K}[Y](x) = \operatorname{co} \mathcal{L}(x).$$

We want to prove that

$$\underbrace{\{\lim X_1(y_k) | y_k \to x_1\} \times \{\lim X_2(z_k) | z_k \to x_2\}}_{\mathcal{H}(x)} \subseteq \mathcal{L}(x)$$

where y_k and z_k are any sequences converging, respectively, to x_1 and x_2. To this aim, for each $x \in \mathbb{R}^{n_1} \times \mathbb{R}^{n_2}$, pick $w \in \mathcal{H}(x)$. By definition, there exist two sequences y_k, z_k converging, respectively, to x_1, x_2, such that

$$w = (\lim X_1(y_k), \lim X_2(z_k)).$$

Define the sequence $x_k = (y_x, z_k)$, and notice that $x_k \to x$. Therefore, since $w = \lim Y(x_k)$, it follows that $w \in \mathcal{L}(x)$. Thus, since the latter construction holds for every $w \in \mathcal{H}(x)$, it follows that for each $x \in \mathbb{R}^{n_1} \times \mathbb{R}^{n_2}$

$$\mathcal{H}(x) \subseteq \mathcal{L}(x).$$

Now we want to prove the complementary inclusion. To this end, for each $x \in \mathbb{R}^{n_1} \times \mathbb{R}^{n_2}$, pick $w \in \mathcal{L}(x)$. Then, by definition, there exists a sequence x_k converging to x, such that $w = \lim Y(x_k)$. Split such a sequence with respect to its components, *i.e.*, $x_k = (y_k, z_k)$. By the definition of Y, if follows that

$$w = (\lim X_1(y_k), \lim X_2(z_k))$$

that is $w \in \mathcal{H}(x)$. Thus, for each $x \in \mathbb{R}^{n_1} \times \mathbb{R}^{n_2}$, $\mathcal{L}(x) \subseteq \mathcal{H}(x)$. The two shown inclusions yield, for each $x \in \mathbb{R}^{n_1} \times \mathbb{R}^{n_2}$,

$$\mathcal{H}(x) = \mathcal{L}(x).$$

To conclude the proof, notice that by taking the convex-hull of both sides of the latter expression gives[4]

$$\begin{aligned} \operatorname{co} \mathcal{L}(x) &= \operatorname{co} \left(\{\lim X_1(y_k) | y_k \to x_1\} \times \{\lim X_2(z_k) | z_k \to x_2\}\right) \\ &= \operatorname{co} \{\lim X_1(y_k) | y_k \to x_1\} \times \operatorname{co} \{\lim X_2(z_k) | z_k \to x_2\} = \mathcal{K}[X_1](x_1) \times \mathcal{K}[X_2](x_2). \end{aligned}$$

[4]We used the following property. Let, for $i = 1, 2, \dots, s$, $S_i \subset \mathbb{R}^{n_i}$ given sets, then $\operatorname{co} \bigtimes_{i=1}^{s} S_i = \bigtimes_{i=1}^{s} \operatorname{co} S_i$; see [13].

■

Remark 1.4. The above result decreases the conservatism of [97, Theorem 1 (3)] for the special class of functions considered. Notice that, whenever for each $i = 1, 2, \dots, s$ $n_i = 1$, the above theorem specializes to the case of decentralized functions.

By using the rules illustrated in the above result, we reconsider Example 1.3 to investigate the existence of Krasovskii solutions φ, with $\varphi(0) = 0$.

Example 1.4. Consider the quantized closed-loop system given in (1.9). Notice that, since the function $\operatorname{sign}(\cdot)$ is locally bounded, (in fact bounded), according to Theorem 1.1, at least for small enough T, there exists a Krasovskii solution to (1.9) for every $x_0 \in \mathbb{R}$. To determine such a solution, one needs to first determine the Krasovskii regularization of $-\operatorname{sign}(x)$. In particular, as for every $x \neq 0$, $\operatorname{sign}(x)$ is continuous, by the items (i) and (iii) of Proposition 1.1, and via the expression given in (1.10), one gets

$$\mathcal{K}[-\operatorname{sign}](x) = \begin{cases} -1 & x > 0 \\ 1 & x < 0 \\ [-1, 1] & x = 0. \end{cases}$$

Differently from Example 1.3, the zero function is a (the unique) Krasovskii solution to (1.9) on any interval of $\mathbb{R}_{\geq 0}$, and obviously $\varphi(0) = 0$. The main difference with respect to Example 1.3 consists of having enabled solutions starting from the origin to be constant.

At this stage, it should be clear that differential inclusions play a key role in this dissertation. In particular, let us consider the following differential inclusion

$$\dot{x} \in F(x) \tag{1.13}$$

where $x \in \mathbb{R}^\ell$, and $F(x) \colon \mathbb{R}^\ell \rightrightarrows \mathbb{R}^\ell$. For such a differential inclusion, let us consider the notion of solution given next.

Definition 1.5. Let $\mathbb{I} \subset \mathbb{R}_{\geq 0}$ be an interval containing 0. The function $\varphi \colon \mathbb{I} \to \mathbb{R}^n$ is a solution to (1.13) if φ is absolutely continuous on $\mathbb{I}$, and

$$\dot{\varphi}(t) \in F(\varphi(t)) \qquad \text{for almost all} \quad t \in \mathbb{I}.$$

The above definition allows to consider Krasovskii solutions to a given differential equation as the solutions to a certain differential inclusion. Therefore, in the sequel, for the sake of generality, results, definitions and properties will be stated for general differential inclusions as (1.13).

Concerning solutions to (1.13), in this dissertation, we consider the following notions.

Definition 1.6 (Maximal solution [56])**.** Let φ be a solution to (1.13). Then φ is said to be maximal if there does not exist any other solution ψ such that $\operatorname{dom} \phi$ is a proper subset of $\operatorname{dom} \psi$ and $\varphi(t) = \psi(t)$ for every $t \in \operatorname{dom} \varphi$.

Definition 1.7 (Complete solution [56])**.** Let φ be a solution to (1.13). Then φ is said to be complete if $\sup \operatorname{dom} \varphi = \infty$.

Remark 1.5. Clearly, every complete solution is maximal but the converse is in general not true.

1.2.2 About Numerical Simulations of Krasovskii Solutions

To overcome the issues about the existence of solutions to (1.7), we addressed the study of such a system, by means of the notion of Krasovskii solution. The adoption of this notion perfectly fits in control problems. On the other hand, when one is interested in the numerical simulation of (1.7), the question that naturally arises is how to integrate (1.7) to somehow recover the behaviors captured by the notion of Krasovskii solution.

For this purpose, we need to introduce the notion of δ-polygonal approximation and Euler solution, which are both given in [27].

Definition 1.8 (δ-Polygonal approximation)**.** Consider system (1.7). Given $x_0 \in \mathbb{R}^n$ and $T > 0$, consider the following construction

- Fix an arbitrary partition of the interval $[0, T]$, $0 < t_1 < t_2 < \cdots < t_N$, with $t_N = T$ and $\max\limits_{k \in \{0,1,\ldots,N-1\}} \{t_{k+1} - t_k\} \leq \delta$.
- Compute $x_{k+1} = x_k + (t_{k+1} - t_k)X(x_k)$, for $k = 0, \ldots, N-1$ and $x(0) = x_0$.
- Build the piecewise affine function $\varphi_\delta(t_k)$ such that $\varphi_\delta(t_k) = x_k$ for $k = 0, 1, \ldots, N-1$.

The function $\varphi_\delta(t)$ is said to be a δ-polygonal approximation for (1.7).

Definition 1.9. A function $\varphi_{\mathcal{E}}(t)$ is said to be an Euler solution to (1.7) if it is the uniform limit for $\delta \to 0$ of a polygonal approximation $\varphi_\delta(t)$ obtained by some partition of the interval $[0, T]$, and for some $x_0 \in \mathbb{R}^n$.

The interest in considering Euler solutions stems from the fact that, as proven in [16], Euler solutions are Krasovskii solutions. In particular, notice that, among all the possible polygonal approximations one can consider, the simplest and straightforwardly attainable through a numerical procedure arises from selecting a uniform partitioning of the time interval $[0, T]$. Namely, let N be an arbitrarily positive integer, fix $\delta = \frac{T}{N}$, set $t_0 = 0$ and for $k = 0, 1, 2, \ldots, N-1$, select $t_{k+1} = t_k + \frac{T}{N}$. Thus, the sequence of polygonal approximations $\{x_{\frac{T}{N}}\}_{N=1}^{\infty}$, if converges uniformly, has as a limit a Krasovskii solution to (1.7). Therefore, for N sufficiently large, the function $x_{\frac{T}{N}}$ can represent a good approximation of a Krasovskii solution to the considered system. This aspect is illustrated in the following example.

Example 1.5. Consider again the system analyzed in Example 1.3, and recall that for such a system, there exists only a maximal Krasovskii solution φ, with $\varphi(0)$, *i.e.*, the null solution. Then, in this case, for every compact interval $[0, T]$, whatever is the partition used to determine δ-polygonal approximations to (1.3), as δ approaches zero, if the family of functions φ_δ converges uniformly on the interval $[0, T]$, its (uniform) limit is the identically zero function on the interval $[0,T]$, that is

$$\lim_{\delta\to 0}\sup_{t\in[0,T]}|\varphi_\delta(t)| = 0.$$

In particular, this fact can be shown numerically in this case by considering a uniform partitioning. Figure 1.2 shows the value of $\sup\limits_{t\in[0,10]}|\varphi_{\frac{T}{N}}(t)|$ versus N. As N approaches infinity (δ approaches zero), $\sup\limits_{t\in[0,10]}|\varphi_{\frac{T}{N}}(t)|$ approaches zero, meaning that $\varphi_{\frac{T}{N}}$ uniformly approaches the zero function on $[0,10]$. Figure 1.3 depicts some δ-polygonal approximations obtained

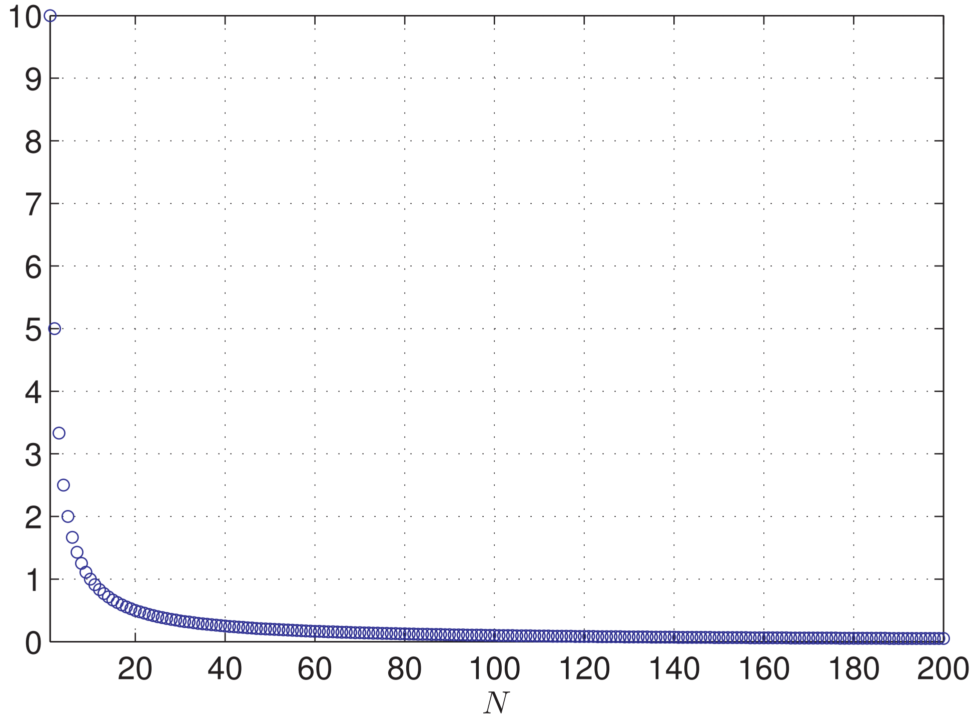


Figure 1.2: $\sup_{s\in[0,10]}|\varphi(s)|$ versus N, for a uniform partitioning.

for different uniform partitioning of the interval $[0,5]$. Figure 1.3 shows that as N increases the resulting δ-polygonal approximation approaches the null solution.

The above example shows that the notion of Euler solution and the fact that Euler solutions are Krasovskii solutions provides some insights on how discontinuous systems could be simulated to capture the peculiar behaviors of Krasovskii solutions. However, following this approach based on Euler first order integration entails two main problems. On the one hand, given $x_0 \in \mathbb{R}^n$, there may exist multiple Krasovskii solutions $\varphi_1, \varphi_2, \dots, \varphi_s$, with $\varphi_1(0) = \varphi_2(0) = \cdots = \varphi_s(0) = x_0$, and some of them may not be Euler solutions. For instance, consider [23, Example 1], for which $X(x) = \frac{3}{2}x^{1/3}$, $T = 1$, and $x_0 = 0$. In this case, it can be shown that $\varphi_1(t) = t^{3/2}$, $\varphi_2(t) = t^{-3/2}$, and $\varphi_3(t) = 0$ are Carathéodory solutions (then obviously Krasovskii solutions) to the considered system with $\varphi_1(0) = \varphi_2(0) = \varphi_3(0) = 0$, while the only Euler solution is $\varphi_{\mathcal{E}}(t) = 0$, despite the continuity of the function X. On the other hand, establishing if the considered sequence of polygonal approximations uniformly converges whenever N approaches infinity could be nontrivial. Therefore, this aspect is still

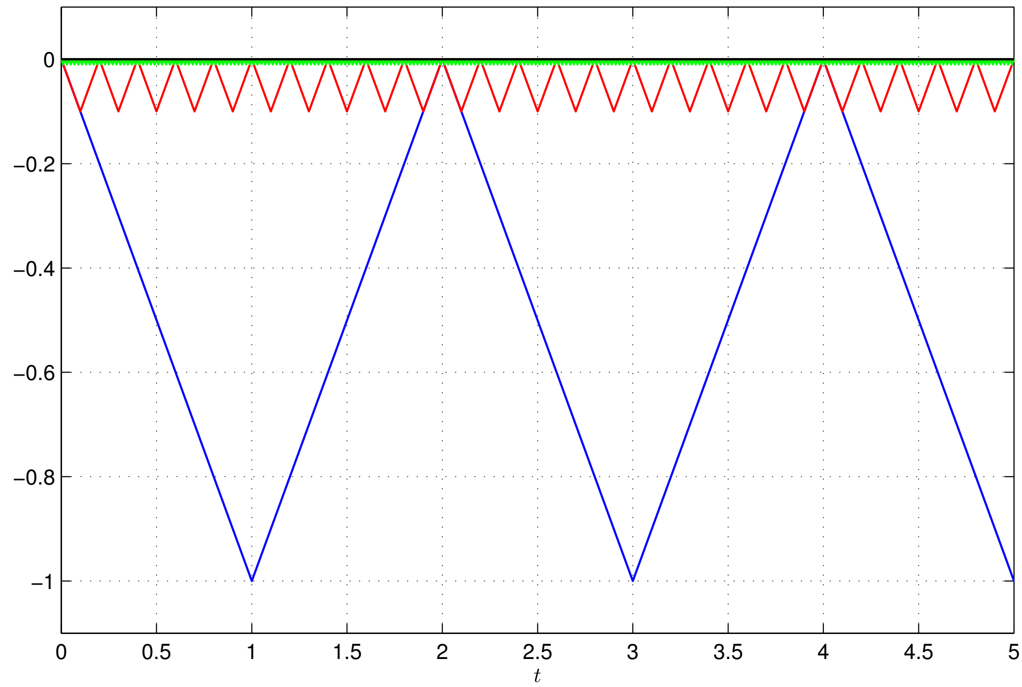


Figure 1.3: Some δ-polygonal approximations ($N = 10$ blue, $N = 100$ red, $N = 1000$ green)

worth of further investigations.

1.3 Uniform Quantized Linear Control Systems

1.3.1 The Class of Systems Under Study

In this dissertation, we focus on plants whose dynamics are linear, that is dynamical systems in the following form

$$\begin{cases} \dot{x} = Ax + Bu \\ y = Cx \end{cases} \tag{1.14}$$

where $A \in \mathbb{R}^{n\times n}$, $B \in \mathbb{R}^{n\times m}$, and $C \in \mathbb{R}^{p\times n}$. For such a class of plants, the following standing assumptions will be considered in the sequel.

Assumption 1.1 (Standing assumption)**.** The matrix A is not Hurwitz. $\triangle$

Assumption 1.2 (Standing assumption)**.** The pair (A, B) is stabilizable, and the pair (A, C) is detectable. $\triangle$

Assumption 1.1 allows to exclude the trivial case of open-loop stable plants. Whereas, Assumption 1.2 ensures that a linear stabilizing controller exists for the considered plant, assumption that will play a fundamental role in our approach.

The interest in considering such a class of systems is twofold. On the one hand, many real plants can be approximately modeled through a linear model, at least around an equilibrium point. On the other hand, by considering linear plants, constructive methodologies can be proposed. Namely, building on theoretical conditions, numerical algorithms for the solution to the analyzed problems can be derived.

In this particular case, being the dynamics of the plant linear, we reasonably consider also linear controllers. Therefore, by specializing the various situations presented earlier to the case of linear plant and linear controllers, we obtain the following models for the closed-loop system.

Linear static state feedback controller.

$$\begin{cases} \dot{x} = Ax + Bu \\ u = \mathrm{q}_u(u_c) \\ u_c = K\,\mathrm{q}_y(x) \end{cases} \tag{1.15}$$

where $K \in \mathbb{R}^{m \times n}$ is the controller gain.

Linear dynamic output feedback controller.

$$\begin{cases} \dot{x} = Ax + Bu \\ \dot{x}_c = A_c x_c + B_c y_m \\ u_c = C_c x_c + D_c y_m \\ u = \mathrm{q}_u(u_c) \\ y_m = \mathrm{q}_y(Cx) \end{cases} \tag{1.16}$$

where $x_c \in \mathbb{R}^{n_c}$ is the controller state, and $A_c \in \mathbb{R}^{n_c \times n_c}, B_c \in \mathbb{R}^{n_c \times p}, C_c \in \mathbb{R}^{m \times n_c}, D_c \in \mathbb{R}^{m \times p}$ are the matrices defining the controller model.

1.3.2 The Uniform Quantizer

In this dissertation, we focus on the uniform quantizer $\mathrm{q}\colon \mathbb{R} \to \Delta\mathbb{Z}$ defined as follows,

$$\mathrm{q}(u) := \Delta \,\mathrm{sign}(u) \left\lfloor \frac{|u|}{\Delta} \right\rfloor \tag{1.17}$$

where Δ is a positive given real scalar characterizing the quantization error bound, *i.e.*, for every u, $|\,\mathrm{q}(u) - u| \leq \Delta$; see Figure 1.4. Whenever, $u \in \mathbb{R}^\ell$, with $\ell > 1$, then

$$\mathrm{q}(u) := (\mathrm{q}(u_1), \mathrm{q}(u_2) \ldots, \mathrm{q}(u_\ell)).$$

Remark 1.6. Observe that the quantizer we consider in this dissertation, due to the larger dead-zone around the origin with respect to a standard quantizer, it is genuinely uniform only when restricted to $\mathbb{R}_{\geq 0}$. The choice of this quantizer stems from having for a given Δ a quantizer as coarse as possible. Indeed, the standard uniform quantizer adopted, *e.g.*, in

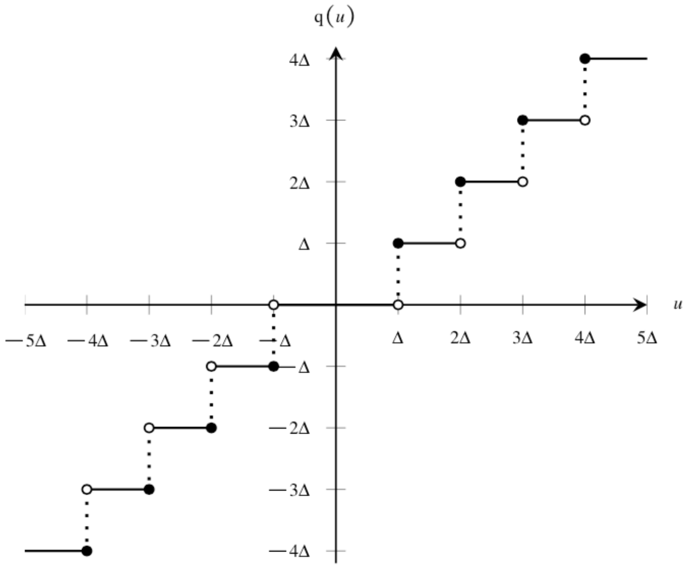


Figure 1.4: The uniform quantizer

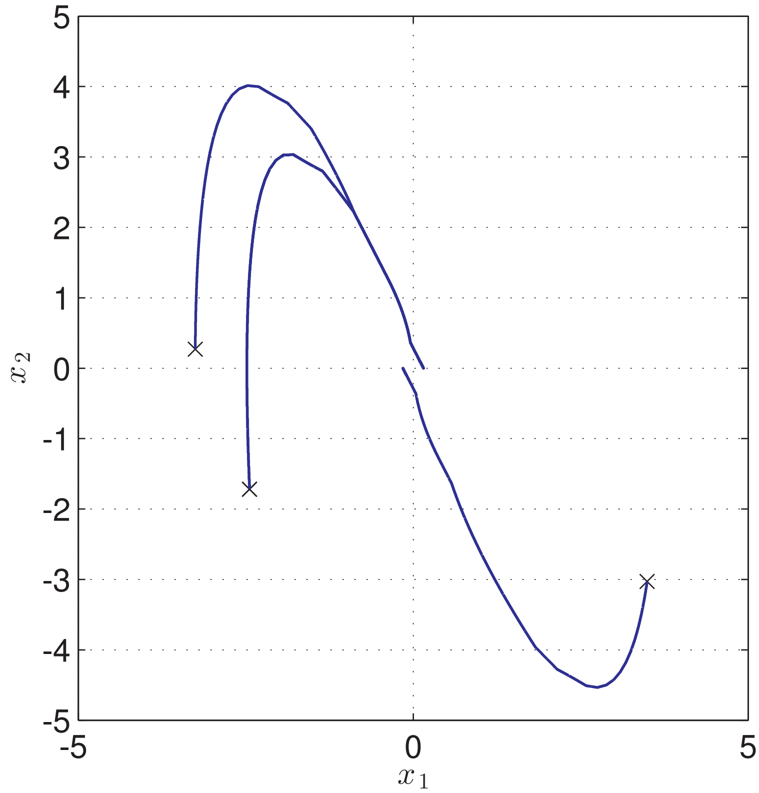
[22], for a given $\Delta > 0$ induces a quantization error bounded by $\frac{\Delta}{2}$. A quantizer similar to (1.17) is considered in [85], although we slightly modified such a map to avoid having a discontinuity at the origin. That said, since the quantizer we consider entails the same bound on the quantization error as the in the case of the uniform quantizer in [85], with a slight abuse of notation, we denote (1.17) uniform quantizer. We would like to emphasize that all the results presented within this dissertation can be easily extended to encompass the standard uniform quantizer used, *e.g.*, in [22].

Notice that, since $\mathrm{q}(0) = 0$, and the plant and the controller dynamics are homogeneous (in fact they are linear), both for (1.15) and (1.16), the origin is an equilibrium point for the closed-loop system. Assume that the origin is also globally asymptotically stable for the quantization free closed-loop system, one may wonder whether the same property still holds for systems (1.15) and (1.16). The following examples show that, in general, the answer to this question is negative.

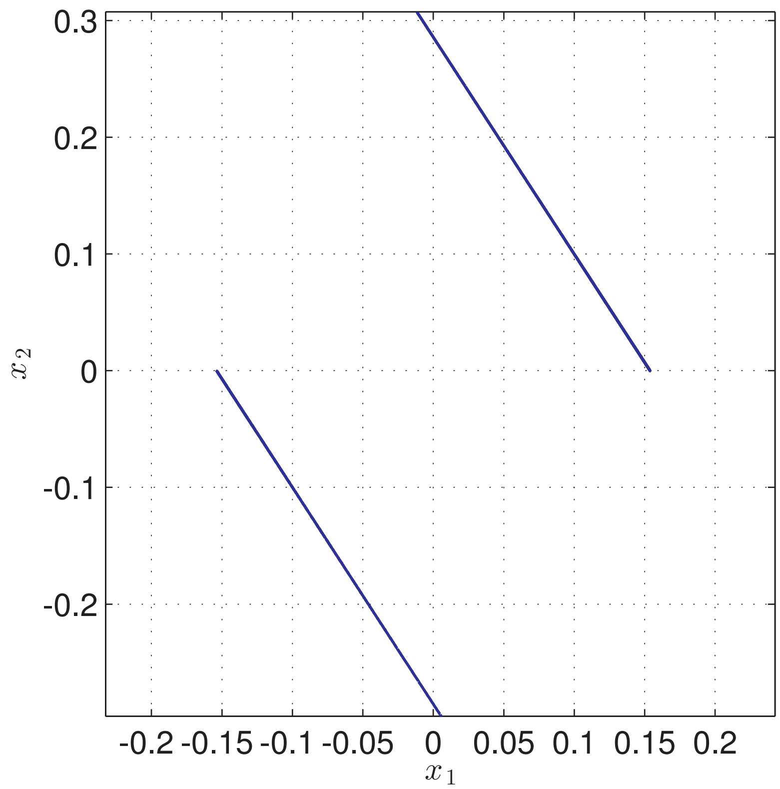
Example 1.6 (Isolated equilibria)**.** Consider the quantized input version of the balancing pointer from [69].

$$\begin{cases} \dot{x} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} x + \begin{bmatrix} 0 \\ -1 \end{bmatrix} u \\ u = \mathrm{q}(u_c) \end{cases}$$

Suppose that the plant is controlled via a static state feedback controller $u_c = Kx$, with $K = \begin{bmatrix} 13 & 7 \end{bmatrix}$, and $\mathrm{q}(\cdot)$ is the uniform quantizer with $\Delta = 2$. Notice that, whenever the plant actuator is not quantized, the origin of the closed-loop system is globally asymptotically stable, as $\mathrm{spec}(A+BK) = \{-3, -4\}$. In Figure 1.5 some closed-loop trajectories are shown. Simulations show that the closed-loop system trajectories approach two isolated equilibrium point. Therefore, the origin is no longer globally asymptotically stable.



(a) Closed-loop trajectories



(b) A close-up showing the trajectories converging toward the two equilibria

Figure 1.5: Quantized control system manifesting isolated equilibria.

Example 1.7 (Limit-cycles)**.** Consider again the balancing pointer plant described in the above example, and assume that the measured state is quantized via a uniform quantizer (1.17) with $\Delta = 0.5$. Suppose that the plant is controlled via the same static state feedback controller given in Example 1.6, *i.e.*, $u_c = K\,\mathrm{q}(x)$.

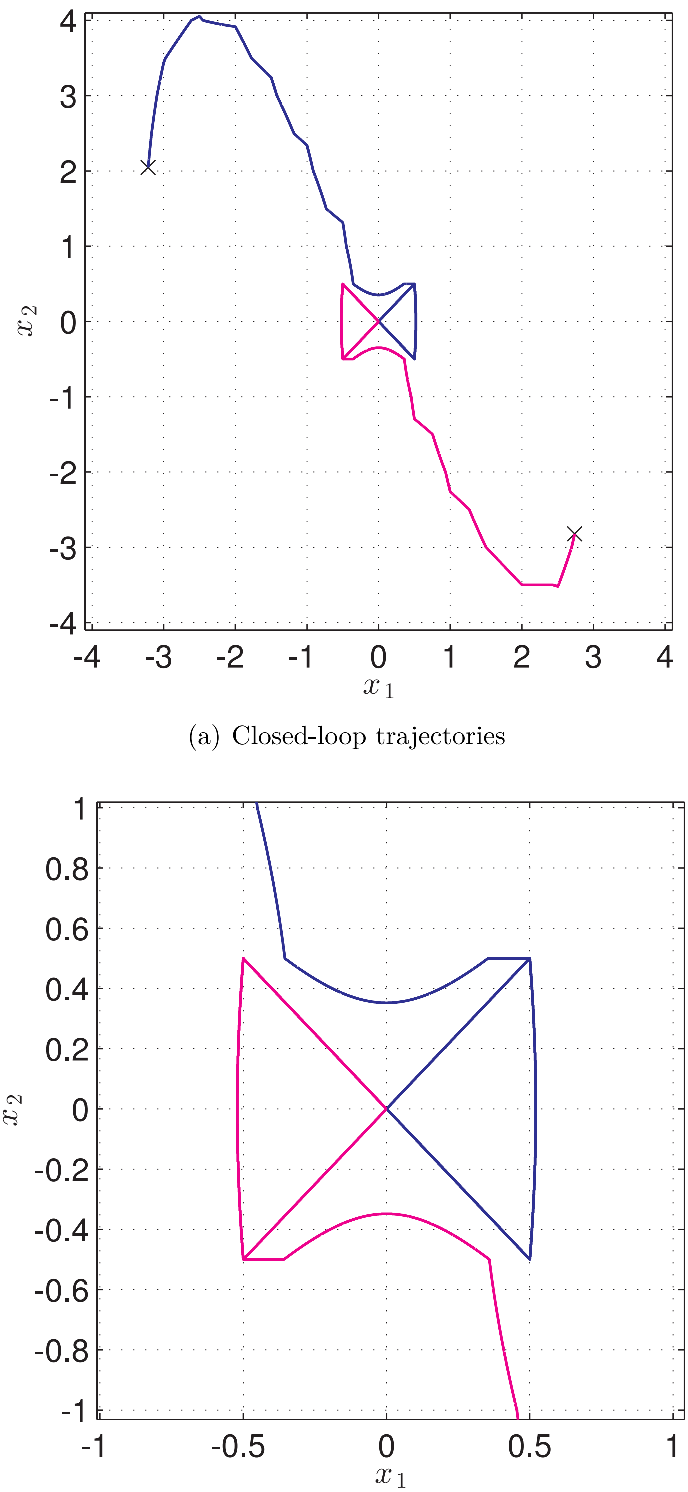
$$\begin{cases} \dot{x} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} x + \begin{bmatrix} 0 \\ -1 \end{bmatrix} u \\ u = u_c \end{cases}$$

In Figure 1.6, some closed-loop trajectories are shown. Simulations show that the closed-loop system trajectories approach a limit cycle, implying that the origin is not globally asymptotically stable.

The two above examples show that, in general, the asymptotic stability properties of the quantization free closed-loop system are destroyed by quantization. This phenomenon is well established in the literature; see, *e.g.*, [22, 84, 117]. In particular, as far as concerns (1.17), due to finite precision near the origin, such a quantizer induces in both (1.15) and (1.16) a region of the state space wherein the control system runs in open loop. This implies that if the origin of the open-loop plant is not asymptotically state, so is the origin of the closed-loop system. For instance, consider system (1.16), and suppose that the origin of the open-loop plant is not asymptotically stable. Let q_u and q_y defined as in (1.17), with respectively Δ_u and Δ_y. Pick $x_c = 0$, and x_0 such that $|Cx_0| \prec \Delta_y$. Now, let φ be a maximal solution to $\dot{x} = Ax$, with $\varphi(0) = x_0$. Due to linearity, there exists a strictly positive T, such that $|C\varphi(t)| \prec \Delta_y$ for each $t \in [0, T]$. Thus, $(\varphi(t), 0)$ is a solution to (1.16) on the interval $[0, T]$. Since this construction can be repeated for any x_0 such that $|Cx_0| \prec \Delta_y$, and the origin of the open-loop plant is not asymptotically stable by hypothesis, so is the origin of (1.16). Basically, sensor quantization induces a lack of the feedback action in a polyhedral region containing the origin, preventing from achieving closed loop asymptotic stability for the origin. Similar arguments show that actuator quantization induces the same kind of behaviors, while analogous considerations hold also for the simpler case of the static state feedback control system (1.15).

1.4 Stability Notion and Preliminaries Results

The facts illustrated above, also via Example 1.6 and Example 1.7, underline that, in general, requiring the origin of the closed-loop system (1.15) or (1.16) to be asymptotically stable is in general impossible. In fact, quantized dynamical systems may manifest complex behaviors, whose precise characterization, unless in particular cases, is far from trivial. On the other hand, as shown in [84, 117], and qualitatively illustrated in Example 1.6 and Example 1.7, under suitable conditions, the closed-loop system trajectories are bounded and converge into a compact and invariant set $\mathcal{A}$ containing the origin, (such a set can contain limit cycles, equilibrium point etc.). Loosely speaking, the set $\mathcal{A}$ gives an outer approximation, near the



(a) Closed-loop trajectories

(b) A close-up showing the trajectories converging toward two limit-cycles

Figure 1.6: Quantized control system manifesting limit-cycles.

origin, of the real behavior of the closed-loop system. In particular, the determination of the set $\mathcal{A}$ enables to define a bounded region having two relevant properties: (1) Closed-loop solutions starting inside $\mathcal{A}$ remain definitely confined in such a set, (2) closed-loop solutions starting outside $\mathcal{A}$ approach such a set. That said, it appears likewise interesting to investigate on what happens when the closed-loop system is initialized "near" such a set. From a technical point of view, this fact prompts to seek for conditions guaranteeing the asymptotic stability of a compact set containing the origin.

In particular, in this dissertation, for a general differential inclusion as (1.13), we consider the following notion of (uniform) global stability for a closed set $\mathcal{A} \subset \mathbb{R}^\ell$, given in [124]. Such a definition uses distance to closed set, and class $\mathcal{K}$ functions, which are given next.

Definition 1.10 (Distance to a closed set [56])**.** Given a vector $x \in \mathbb{R}^n$, and a closed set $\mathcal{A}$, the distance of x from $\mathcal{A}$ is denoted $|x|_\mathcal{A}$ and is defined by $|x|_\mathcal{A} = \inf_{y\in\mathcal{A}} \|x - y\|$.

Remark 1.7. Notice that, given a closed set $\mathcal{A} \subset \mathbb{R}^n$ and a positive real scalar δ, the set of the points $x \in \mathbb{R}^n$ with $|x|_\mathcal{A} \leq \delta$ coincides with the set $\mathcal{A} + \delta\mathbb{B}$. Such a writing will be largely used throughout this dissertation.

Definition 1.11 (Class $\mathcal{K}_\infty$ functions [76])**.** A function $\alpha\colon \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$, is a class $\mathcal{K}_\infty$ if α is zero at zero, continuous, strictly increasing, and unbounded.

The definition of uniform global asymptotic stability of a closed-set is as follows.

Definition 1.12 (Uniform global asymptotic stability)**.** Let $\mathcal{A} \subset \mathbb{R}^n$ be closed. The set $\mathcal{A}$ is

- uniformly globally stable for (1.13), if there exists a class $\mathcal{K}_\infty$ function α, such that every solution φ to (1.13) satisfies $|\varphi(t)|_\mathcal{A} \leq \alpha(|\varphi(0)|_\mathcal{A})$ for every $t \in \operatorname{dom}\varphi$
- uniformly globally attractive for (1.13), if every maximal solution to (1.13) is complete, and for every $\varepsilon > 0$ and $\mu > 0$ there exists $T > 0$, such that for any solution φ to (1.13) with $|\varphi(0)|_\mathcal{A} \leq \mu$, $t \geq T$ implies $|\varphi(t)|_\mathcal{A} \leq \varepsilon$
- uniformly globally asymptotically stable (*UGAS*) for (1.13), if it is uniformly globally stable and uniformly globally attractive

The uniformity requirement considered in the above notion of stability implies that whenever the distance of the initial condition $\varphi(0)$ from the set $\mathcal{A}$ approaches zero, so does the distance of the issuing solution $\varphi(t)$ for each $t \in \operatorname{dom}\varphi$. The uniformity requirement considered in the attractivity property implies instead that the convergence rate of the solutions' distance from the set $\mathcal{A}$ is uniform with respect to the initial condition's distance. Although the uniformity requirements considered in the above definition gives rise to stronger notions of stability than the one usually considered, it turns out that for the class of systems and problems addressed in this dissertation, the uniformity requirement is without loss of generality. This aspect will be clarified through the results given in the sequel.

For the special case of compact sets, let us consider the following result which essentially derives from the combined application of [56, Proposition 7.5.] and [124, Proposition 3]. The derivation of such a result uses the definition of strong forward invariance of a closed set for a differential inclusion, given *e.g.*, in [26] and reported below, and general definitions concerning set-valued mappings that are reported in Appendix D.

Definition 1.13. Let $\mathcal{A} \subset \mathbb{R}^n$ be closed. The set $\mathcal{A}$ is strongly forward invariant for (1.13) if every maximal solution to (1.13) is complete, and $\varphi(0) \in \mathcal{A}$ implies $\operatorname{rge}\varphi \subset \mathcal{A}$.

Now we are in position to state the mentioned result.

Proposition 1.4. *Consider the differential inclusion in* (1.13)*, i.e.,*

$$\dot{x} \in F(x) \qquad x \in \mathbb{R}^n, F\colon \mathbb{R}^n \rightrightarrows \mathbb{R}^n.$$

Let $\mathcal{A} \subset \mathbb{R}^\ell$ be compact, strongly forward invariant and uniformly globally attractive for (1.13)*. Let F be outer semicontinuous, locally bounded,* $\operatorname{dom} F = \mathbb{R}^n$*, and such that for each $x \in \mathbb{R}^n$ $F(x)$ is convex. Then, the set $\mathcal{A}$ is UGAS for* (1.13)*.* □

The proof of the above result uses class $\mathcal{KL}$ functions.

Definition 1.14 (Class $\mathcal{KL}$ functions [76])**.** A function $\beta\colon \mathbb{R}_{\geq 0} \times \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$, is a class $\mathcal{KL}$ function, if it is nondecreasing in its first argument, nonincreasing in its second argument, and

$$\lim_{s \to 0^+} \beta(s,t) = \lim_{t \to +\infty} \beta(s,t) = 0.$$

Then, the proof of the above result is as follows.

Proof of Proposition 1.4. Due to the properties required for F in the statement of the above result, since $\mathcal{A}$ is compact, strongly forward invariant, and uniformly globally attractive for (1.13), thanks to [56, Proposition 7.5.] it follows that $\mathcal{A}$ is stable[5] for (1.13). Moreover, due to the properties required for F, by the virtue of [124, Proposition 3] it follows that there exists a class-$\mathcal{KL}$ function β, such that for every maximal solution φ to (1.13), one has for every $t \in \mathbb{R}_{\geq 0}$,

$$|\varphi(t)|_{\mathcal{A}} \leq \beta(|\varphi(0)|_{\mathcal{A}}, t)$$

which in turn, due to [124, Proposition 1], implies that $\mathcal{A}$ is UGAS for (1.13), and this finishes the proof. ■

Notice that the above result plays a fundamental role in establishing sufficient conditions to ensure UGAS of a certain compact set containing the origin. Indeed, as previously illustrated in this chapter, the requirements on the right-hand side set-valued mapping $F(x)$ needed for the applicability of Proposition 1.4 are obviously verified whenever $F(x)$ arises from the Krasovskii regularization of a locally bounded function, which is the case in both (1.15) and (1.16).

Remark 1.8. UGAS of a compact set $\mathcal{A}$ for (1.13) ensures that every maximal solution to (1.13) is bounded. To see this, it suffices to observe that, being $\mathcal{A}$ compact, for a large enough $\delta > 0$, one has $\mathcal{A} \subset \delta\mathbb{B}$. Thus, since for every $x \in \mathbb{R}^n$, $|x|_{\delta\mathbb{B}} \leq |x|_{\mathcal{A}}$, and $\|x\| \leq |x|_{\delta\mathbb{B}} + \delta$. Finally, boundedness of maximal solutions to (1.13) can be readily established by combining the latter relations with the bounds issued from UGAS.

Before concluding this chapter, let us consider the following result, which will be exploited in the sequel.

Proposition 1.5. *Consider* (1.13) *and assume that F is outer semicontinuous, locally bounded, convex valued, and* $\operatorname{dom} F = \mathbb{R}^n$*. Assume that there exists a continuously dif-*

[5]See, *e.g.*, [124, Proposition 3] for a standard definition of $\varepsilon - \delta$ stability of a compact set.

ferentiable function $V\colon \mathbb{R}^n \to \mathbb{R}$ *such that*

$$V(x) > 0 \qquad \forall x \neq 0 \tag{1.18}$$

$$\lim_{\|x\|\to\infty} V(x) = \infty \tag{1.19}$$

and two positive real scalars ρ*, and* α *such that*

$$\langle \nabla V(x), f\rangle \leq -\rho V(x) \qquad \forall x \in \mathcal{L}^+_\alpha(V), f \in F(x) \tag{1.20}$$

where $\mathcal{L}^+_\alpha(V) := \{x \in \mathbb{R}^n \colon V(x) \geq \alpha\}$*. Then, the set* $\mathcal{A} := \mathbb{R}^n \setminus \mathrm{Int}\mathcal{L}^+_\alpha(V)$ *is UGAS for* (1.13)*.*

The proof of the above result rests on the following lemma.

Lemma 1.1. Let $\mathcal{A} \subset \mathbb{R}^n$ be compact. If there exists a continuous function $\Upsilon\colon \mathbb{R}^n \to \mathbb{R}_{\geq 0}$ such that for each $\mu > 0$, every maximal solution φ to (1.13) with $\varphi(0) \in \mathcal{A} + \mu\mathbb{B}$ is complete, and $t \geq \Upsilon(x_0)$ implies $\varphi(t) \in \mathcal{A}$. Then, $\mathcal{A}$ is globally uniformly attractive for (1.13). □

Proof. The proof is straightforward. In particular, let $\mu > 0$ define

$$\xi = \max_{x \in \mathcal{A}+\mu\mathbb{B}} \Upsilon(x)$$

and observe that being Υ continuous and $\mathcal{A}$ compact, ξ is well defined. To conclude, notice that for each maximal solution φ to (1.13) with $\varphi(0) \in \mathcal{A} + \mu\mathbb{B}$, one has that $t \geq \xi$ implies $\varphi(t) \in \mathcal{A}$ and this concludes the proof. ■

Remark 1.9. The main feature of the above result consists of establishing uniform attractivity via finite time convergence, assuming continuous dependence of the convergence time on the initial condition. Specifically, the continuity requirement allows to establish uniformity with respect to the initial condition.

Now we are in position to show the proof of Proposition 1.5.

Proof of Proposition 1.5. First observe that since V is radially unbounded, $\mathcal{A}$ is compact. To prove that the set $\mathcal{A}$ is UGAS, we firstly show that $\mathcal{A}$ is strongly forward invariant for (1.13) and that each maximal solution to (1.13) is complete.

Concerning strongly forward invariance, since $\mathcal{A}$ is compact, thanks to the properties required for F, from [56, Proposition 6.10.], it suffices to show that each maximal solution starting inside $\mathcal{A}$ cannot leave such a set, *i.e.*, completeness of such solutions automatically holds. By contradiction, assume that there exists a maximal solution φ starting from $\mathcal{A}$ that eventually leaves such a set. Then, there exists $\tau \in \mathrm{dom}\,\varphi$ such that $\varphi(\tau) \notin \mathcal{A}$, that is

$$V(\varphi(\tau)) > \alpha.$$

Thus, since the function $V \circ \varphi\colon \mathrm{dom}\,\varphi \to \mathbb{R}$ is continuous, there exists $s \in \mathrm{dom}\,\varphi$ such that

$$V(\varphi(s)) = \alpha.$$

Without loss of generality[6], assume that for each $t \in (s, \tau]$, $\varphi(t) \notin \mathcal{A}$. In other words, s is the largest exit time of the solution φ from the set $\mathcal{A}$. From (1.20) thanks to the Grönwall-Bellman lemma, it follows that for every $t \in [s, \tau]$

$$V(\varphi(t)) \leq e^{-\rho(t-s)} V(\varphi(s))$$

then

$$V(\varphi(\tau)) < V(\varphi(s)).$$

However, this contradicts the fact that $\varphi(\tau) \notin \mathcal{A}$, *i.e.*, φ cannot leave the set $\mathcal{A}$. Hence, $\mathcal{A}$ is strongly forward invariant for (1.13).

Concerning completeness of the maximal solutions starting outside $\mathcal{A}$, by retracing the same steps performed above, it can be readily shown that every maximal solution φ to (1.13) and with $\varphi(0) \notin \mathcal{A}$ cannot leave the sublevel set $\mathcal{L}^{-}_{V(\varphi(0))}(V) := \{x \in \mathbb{R}^n \colon V(x) \leq V(\varphi(0))\}$. Hence, since sublevel sets of V are compact, it follows that every maximal solution to (1.13) is bounded. Thus, thanks to [56, Proposition 6.10.], every maximal solution to (1.13) is complete.

Bearing in mind completeness of maximal solutions to (1.13) and strong forward invariance of $\mathcal{A}$, now we conclude the proof of the above result by showing that maximal solutions to (1.13) converge in finite time into $\mathcal{A}$. Pick any maximal solution φ to (1.13), with $\varphi(0) \in \mathrm{Int}\Theta$. Let $\mathcal{T} = \{t \in \mathbb{R}_{\geq 0} \colon \varphi(t) \in \mathcal{A}\}$, since $\mathcal{A}$ is strongly forward invariant, either $\mathcal{T} = \emptyset$ or $\sup \mathcal{T} = \infty$. In other words, if φ eventually enters $\mathcal{A}$, then by strong forward invariance, it cannot leave such a set. By contradiction, let us suppose that $\mathcal{T} = \emptyset$, then for every $t \in \mathbb{R}_{\geq 0}$, $\varphi(t) \notin \mathcal{A}$. Therefore, still from (1.20), it follows that

$$V(\varphi(t)) \leq e^{-\rho t} V(\varphi(0)) \qquad \forall t \in \mathbb{R}_{\geq 0}. \tag{1.21}$$

Pick,

$$t \geq \frac{1}{\rho} \ln \left(V(\varphi(0)) \frac{1}{\alpha} \right)$$

from (1.21) one gets

$$V(\varphi(t)) \leq \alpha$$

that is $\varphi(t) \in \mathcal{A}$, but this contradicts the fact that $\mathcal{T} = \emptyset$. Now, for every $w \in \mathbb{R}^n$, define

$$\Upsilon(w) := \begin{cases} 0 & w \in \mathcal{A} \\ \frac{1}{\rho} \ln \left(V(w) \frac{1}{\alpha} \right) & w \notin \mathcal{A} \end{cases}$$

notice that Υ is continuous on $\mathbb{R}^n$, and that for every maximal solution ϕ to (1.13), $t \geq \Upsilon(\phi(0))$ implies that $\phi(t) \in \mathcal{A}$. Then, since every maximal solution to (1.13) is complete, from Lemma 1.1 it follows that $\mathcal{A}$ is globally uniformly attractive for system (1.13). Now,

[6]This assumption, is discussed in [11] and for self completeness simple arguments justifying such an assumption are given in Appendix A. Notice that, since $\dot{\varphi}(s)$ may not exist, standard arguments revolving of the monotonicity of the function $t \mapsto V \circ \varphi(t)$ cannot be exploited to conclude.

$\mathcal{A}$ is compact, strongly forward invariant, and globally uniformly attractive for (1.13), from Proposition 1.4 it follows that $\mathcal{A}$ is UGAS for (1.13), and this finishes the proof. ■

1.5 Conclusion

In this chapter, we illustrated the quantization phenomena in control systems, with a special attention to uniform quantization and linear control systems. In particular, two main points were addressed. The first pertains to the notion of solution to adopt to deal with quantized control systems. In particular, it was shown that the discontinuity introduced by quantizers may jeopardize the existence of closed-loop solutions. This issue is completely overcame by considering, for the closed-loop system, the notion of solution due to Krasovskii. The other main aspect highlighted in this chapter regards instead the more convenient notion of stability to adopt in dealing with quantized control systems. Indeed, for a general quantized control system, requiring the asymptotic stability of the origin is unattainable. In this setting, it was shown that considering the asymptotic stability of a compact set containing the origin provides a way to guarantee a proper behavior of the closed-loop system, while matching with the nature of considered problem.

This page is intentionally left blank.

2

QUANTIZED LINEAR STATIC STATE FEEDBACK CONTROL

"Research is what I'm doing when I don't know what I'm doing".

– Wernher von Braun

2.1 Introduction

This chapter pertains to quantization in linear static state feedback control schemes. In particular, two cases are considered. In the first one, the plant state is assumed to be fully measurable and the plant actuator uniformly quantized. In the second one, the plant state is assumed to be fully measured via a uniformly quantized sensor. In such two situations, we address both stability analysis and stabilization of the closed-loop system. The approach followed to address the two configurations is essentially the same. Namely, as a first step we provide a general result to characterize the behavior of the closed-loop system, such a result to some extent uses ideas from [84], though adapted to deal with Krasovskii solutions and uniform global asymptotic stability of a certain compact set. Then, by the use of novel sector conditions, a less conservative result, based on the solution to certain matrix inequalities, is proposed. Building on such a result, a complete apparatus revolving on convex optimization is presented to solve both the stability analysis and the stabilization problems, while taking into account optimization aspects. First results concerning the actuator quantization case can be found in [40].

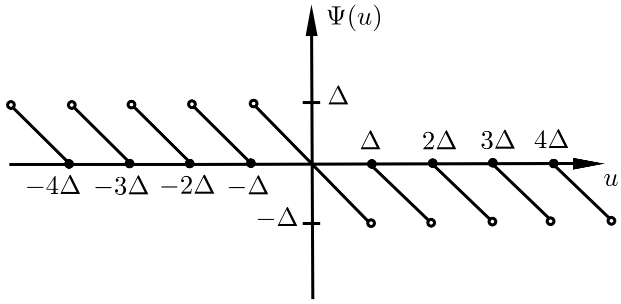


Figure 2.1: The function Ψ, in the scalar case, representing the quantization error.

2.2 Actuator Quantization

2.2.1 Problem Statement and Preliminary Results

Consider the following continuous-time linear system with actuator quantization

$$\begin{cases} \dot{x} = Ax + Bu \\ u = \mathrm{q}(Kx) \end{cases} \tag{2.1}$$

where $x \in \mathbb{R}^n$, $u \in \mathbb{R}^m$, are respectively the state, and the input of the system. A, B, K are real matrices of suitable dimensions, and $\mathrm{q}(\cdot)$ is the uniform quantizer defined in (1.17) having as a quantization error bound $\Delta > 0$. Define the function,

$$\begin{aligned} \Psi\colon \mathbb{R}^m &\to \mathbb{R}^m \\ z &\mapsto \mathrm{q}(z) - z \end{aligned} \tag{2.2}$$

the closed-loop system can be rewritten as

$$\dot{x} = (A + BK)x + B\Psi(Kx). \tag{2.3}$$

The function Ψ represents the quantization error, then according to (1.17), Ψ is bounded. In particular, for every $u \in \mathbb{R}^m$, $\|\Psi(u)\| \leq \sqrt{m}\Delta$; see Figure 2.1. Moreover, since the function Ψ is discontinuous, the right-hand side of (2.3) is a discontinuous function of the state. Thus, for the reasons illustrated in Chapter 1, we focus on Krasovskii solutions to system (2.3). Notice that, in view of the local boundedness of the right-hand side of (2.3), for every $x_0 \in \mathbb{R}^n$, there exists at least a Krasovskii solution φ to (2.3) with $\varphi(0) = x_0$; see Chapter 1. Therefore, by defining

$$\begin{aligned} X\colon \mathbb{R}^n &\to \mathbb{R}^n \\ x &\mapsto (A + BK)x + B\Psi(Kx) \end{aligned} \tag{2.4a}$$

we consider the solutions to the following differential inclusion

$$\dot{x} \in \mathcal{K}[X](x) \tag{2.4b}$$

where $\mathcal{K}[X](x)$ represents the Krasovskii regularization of the function X; see Definition 1.2 on page 14. As pointed out earlier, the presence of the uniform quantizer, due to its deadzone effect, represents a real obstacle to the asymptotic stabilization of the closed-loop system. Namely, one should be aware that if the matrix A is not Hurwitz, then the asymptotic stability of the origin for the closed-loop system (2.4) cannot be achieved via any choice of the gain K. Indeed, for every x belonging to the set $\mathcal{P} := \{x \in \mathbb{R}^n \colon |Kx| \prec \Delta\}$, one has $\Psi(Kx) = -Kx$. Thus, there exists a sufficiently small neighborhood of the origin strictly contained in $\mathcal{P}$, such that for every x the right-hand side of (2.1) coincides with Ax. Namely, the behavior of the closed-loop system around the origin is not influenced by the choice of the gain K. On the other hand, since the function Ψ is bounded, one may expect that, under opportune hypothesis on the quantization free closed-loop system, the solutions to (2.4) manifest some stability properties. A positive answer to this question is given by the following theorem, which uses ideas from [82, Lemma 1].

Theorem 2.1. Let A, B, K be matrices of adequate dimensions, such that $A + BK$ is Hurwitz. Then, there exists a compact set $\mathcal{A} \subset \mathbb{R}^n$, containing the origin, which is UGAS for (2.4).

Proof. Since $A + BK$ is Hurwitz, there exists $P, Q \in \mathcal{S}_+^n$ such that $\operatorname{He}(P(A + BK)) = -Q$. For every $x \in \mathbb{R}^n$, define $\kappa(x) = Kx$. Since the function $x \mapsto (A + BK)x$ is continuous, by Proposition 1.1, for every $x \in \mathbb{R}^n$,

$$\mathcal{K}[X](x) = (A + BK)x + B\mathcal{K}[\Psi \circ \kappa](x).$$

Since Ψ is locally bounded, (in fact bounded), according to [9, Lemma 1] it follows that, for every $x \in \mathbb{R}^n$

$$\mathcal{K}[\Psi \circ \kappa](x) = \operatorname{co}\left\{\lim \Psi(K(x_k)) | x_k \to x\right\}.$$

Then, due to the bound shown earlier on the function Ψ, it turns out that for each $x \in \mathbb{R}^n$

$$\mathcal{K}[\Psi \circ \kappa](x) \subset \mathbb{B}\sqrt{m}\Delta.$$

Therefore, for each $x \in \mathbb{R}^m$, the following inclusion holds:

$$\mathcal{K}[X](x) \subset (A + BK)x + B\mathbb{B}\sqrt{m}\Delta. \tag{2.5}$$

Now, for every $x \in \mathbb{R}^n$, define the function $V(x) = x^\mathsf{T} P x$, and notice that for every $x \in \mathbb{R}^n$, and any $f \in \mathcal{K}[X](x)$

$$\langle \nabla V(x), f \rangle = -x^\mathsf{T} Q x + 2x^\mathsf{T} P B \xi \leq -\lambda_{\min}(Q) x^\mathsf{T} x + 2x^\mathsf{T} P B \xi$$

for some $\xi \in \mathbb{B}\sqrt{m}\Delta$. Let us recall that for every $a, b \in \mathbb{R}^n$ and for every positive scalar ϵ, $2a^\mathsf{T} b \leq \epsilon a^\mathsf{T} a + \frac{1}{\epsilon} b^\mathsf{T} b$. Then, by setting $\epsilon = \frac{1}{2}\lambda_{\min}(Q)$, from the latter inequality one gets

$$\langle \nabla V(x), f \rangle \leq -\frac{1}{2}\lambda_{\min}(Q) x^\mathsf{T} x + \frac{2}{\lambda_{\min}(Q)} \left\| B^\mathsf{T} P^2 B \right\| m\Delta^2 \qquad \forall x \in \mathbb{R}^n, f \in \mathcal{K}[X](x) \tag{2.6}$$

which in turn gives

$$\langle \nabla V(x), f\rangle \leq -\frac{\lambda_{\min}(Q)}{2\lambda_{\max}(P)} V(x) + \frac{2}{\lambda_{\min}(Q)} \left\| B^{\mathsf{T}} P^2 B \right\| m\Delta^2 \qquad \forall x \in \mathbb{R}^n, f \in \mathcal{K}[X](x). \quad (2.7)$$

Pick $\theta \in (0,1)$ and consider the following superlevel set of V

$$\Theta = \left\{ x \in \mathbb{R}^n \colon V(x) \geq \frac{4\lambda_{\max}(P)}{\lambda_{\min}^2(Q)\theta} \left\| B^{\mathsf{T}} P^2 B \right\| \Delta^2 m \right\}$$

and define $\mathcal{A} = \mathbb{R}^n \setminus \operatorname{Int}\Theta$. Moreover, from (2.7)

$$\langle \nabla V(x), f\rangle \leq -\frac{\lambda_{\min}(Q)}{2\lambda_{\max}(P)}(1-\theta)V(x) \qquad \forall x \in \Theta, f \in \mathcal{K}[X](x). \quad (2.8)$$

Then, thanks to Proposition 1.5 it follows that $\mathcal{A}$ is UGAS, completing the proof. ■

Theorem 2.1 shows that if the matrix $A + BK$ is Hurwitz, then there exists a compact set $\mathcal{A}$ containing the origin, which is UGAS for (2.4). Moreover, such a set is a sublevel set of a certain quadratic function. On the one hand, this fact fosters to consider quadratic Lyapunov-like functions to investigate the dynamics of (2.4). This fact essentially arises from the fact that the underlying dynamics of the considered control systems are linear. On the other hand, the characterization of the set $\mathcal{A}$ provided by the above result is quite coarse, and strongly depends on the choice of the matrix Q. It appears obvious that the matrix Q should be selected in a way such that the resulting set $\mathcal{A}$ fits as much as possible the real behavior of the closed-loop system. However, the selection strategy of such a matrix appears unclear. To overcome this problem, we pursue a constructive approach. Namely, first we derive computationally tractable conditions aimed at providing a characterization of the set $\mathcal{A}$. Essentially, through this stage, one obtains a set of conditions whose solution yields the set $\mathcal{A}$. Then, the search of the set $\mathcal{A}$ is done by embedding the obtained conditions into an optimization scheme aimed at shrinking the size of $\mathcal{A}$. The outcome of this approach consists of a systematic procedure able to perform a search of the most convenient set $\mathcal{A}$, starting from the data of the closed-loop system. To operate this approach, we seek for conditions solving the problem formalized as follows.

Problem 2.1. (Stability analysis) Let A, B, K be matrices of adequate dimensions, such that $A + BK$ is Hurwitz. Determine a compact set $\mathcal{A} \subset \mathbb{R}^n$ containing the origin, such that $\mathcal{A}$ is UGAS for system (2.4).

The solution to the above problem is the object of the remainder of this section.

2.2.2 Stability Analysis

As explained earlier, in solving Problem 2.1, we are interested in deriving a set $\mathcal{A}$ fitting as much as possible the real behavior of the closed-loop system. To this end, we want to reduce the conservatism introduced in the proof of Theorem 2.1 to bound the set-valued mapping $\mathcal{K}[\Psi]$. Inspired by the general idea pursued in the literature on nonlinear systems with isolated nonlinearities; see, *e.g.*, [66, 120] and the references therein, we provide some sector conditions providing tighter bounds for the set-valued mapping $\mathcal{K}[\Psi]$. To this aim, consider this first result concerned with the function Ψ.

Lemma 2.1. [38] Let $z \in \mathbb{R}^\ell$, and $S_1, S_2 \in \mathcal{D}_+^\ell$. The following relations hold:

$$\Psi(z)^\mathsf{T} S_1 \Psi(z) - \operatorname{trace}(S_1)\Delta^2 \leq 0 \tag{2.9}$$

$$\Psi(z)^\mathsf{T} S_2(\Psi(z) + z) \leq 0 \tag{2.10}$$

Proof. Let $z = (z_1, z_2, \dots, z_\ell)$. Then, by definition, for each $i \in \{1, 2, \dots, \ell\}$, $|\Psi_i(z)| = |\Psi(z_i)| \leq \Delta$. Now, let $s_1^{(1)}, s_1^{(2)}, \dots, s_1^{(\ell)}$ any strictly positive scalars. One has, for each $i \in \{1, 2, \dots, \ell\}$, $s_1^{(i)}|\Psi_i(z)| \leq s_1^{(i)}\Delta$, then by summing over $i = 1, 2, \dots, \ell$, and by setting

$$S_1 = \operatorname{diag}(s_1^{(1)}, s_1^{(2)}, \dots, s_1^{(\ell)})$$

yields (2.9). To prove (2.10), notice that by definition, for each $i \in \{1, 2, \dots, \ell\}$, $\Psi_i^2(z) + \Psi_i(z)z_{(i)} \leq 0$ (see Figure 2.1). Pick $s_2^{(1)}, s_2^{(2)}, \dots, s_2^{(\ell)}$ any strictly positive scalars. Then, by following the same arguments adopted to show (2.9), and by defining

$$S_2 = \operatorname{diag}(s_2^{(1)}, s_2^{(2)}, \dots, s_2^{(\ell)})$$

yields (2.10), and this concludes the proof. ∎

The above Lemma allows to embed the function Ψ in a certain sector. However, the conditions provided by such a result do not directly apply to the set-valued mapping $\mathcal{K}[\Psi]$, and then further work is needed. On the other hand, let us remark that for every $z \in \mathbb{R}^\ell$ such that $\Psi(z)$ is continuous, as shown in Proposition 1.1, $\mathcal{K}[\Psi](z) = \{\Psi(z)\}$. Then, for such z the conditions provided by Lemma 2.1 are certainly fulfilled. Therefore, the main point to address consists in verifying whether the conditions provided by Lemma 2.1 hold even for the set valued map $\mathcal{K}[\Psi]$ or not. A positive answer to this question is given by the following result.

Lemma 2.2. Let $z \in \mathbb{R}^\ell$, $v \in \mathcal{K}[\Psi](z)$, and $S_1, S_2 \in \mathcal{D}_+^\ell$. Then, the following relations hold:

$$v^\mathsf{T} S_1 v - \operatorname{trace}(S_1)\Delta^2 \leq 0 \tag{2.11}$$

$$v^\mathsf{T}\, S_2\, (v + z) \leq 0 \tag{2.12}$$

Proof. First of all, for each $z \in \mathbb{R}^\ell$, let us define the set

$$\mathcal{L}(z) = \{\lim \Psi(z_k) | z_k \to z\} \subset \mathbb{R}^\ell$$

where z_k is any sequence converging to z. Since Ψ is locally bounded, likewise the proof of [23, Proposition 11], it turns out that for every $z \in \mathbb{R}^\ell$, $\mathcal{K}[\Psi](z) = \operatorname{co}\mathcal{L}(z)$. Now, let us define the following closed set

$$\mathcal{V}_1 = \{v \in \mathbb{R}^\ell \colon v^\mathsf{T} S_1 v - \mathrm{trace}(S_1)\Delta^2 \leq 0\} \subset \mathbb{R}^\ell$$

which, due to S_1 positive definite, is also convex [1]. We want to show that, for $z \in \mathbb{R}^\ell$

$$\mathrm{co}\,\mathcal{L}(z) \subset \mathcal{V}_1. \tag{2.13}$$

To this end, pick $l \in \mathcal{L}(z)$, then, by definition, there exists a sequence $z_k \to z$ such that $l = \lim \Psi(z_k)$. On the other hand, from Lemma 2.1, it turns out that, for every $k \in \mathbb{N}$, and for every diagonal positive definite matrix S_1, one has $\Psi^\mathsf{T}(z_k) S_1 \Psi(z_k) - \mathrm{trace}(S_1)\Delta^2 \leq 0$, which, by taking the limit over k yields $l^\mathsf{T} S_1 l - \mathrm{trace}(S_1)\Delta^2 \leq 0$, that is $l \in \mathcal{V}_1$. Hence,

$$\mathcal{L}(z) \subset \mathcal{V}_1.$$

Thus, since $\mathcal{V}_1$ is convex, taking the convex-hull of both sides of the latter relation establishes (2.13), which in turn gives (2.11).

To show (2.12), we pursue a similar approach. Specifically, for any $z \in \mathbb{R}^\ell$, define the closed set

$$\mathcal{V}_2(z) = \{v \in \mathbb{R}^\ell \colon v^\mathsf{T} S_2 (v + z) \leq 0\} \subset \mathbb{R}^\ell$$

which is convex due to S_2 positive definite. We want to show that $\mathrm{co}\,\mathcal{L}(z) \subset \mathcal{V}_2(z)$. To this end, pick any $l \in \mathcal{L}(z)$, then there exits a sequence $z_k \to z$, such that $l = \lim \Psi(z_k)$. Still, according to Lemma 2.1, for every $k \in \mathbb{N}$, one has $\Psi^\mathsf{T}(z_k) S_2 (\Psi(z_k) + z_k) \leq 0$, then by taking the limit over k, one gets $l^\mathsf{T} S_2 (l + z) \leq 0$, that is $l \in \mathcal{V}_2(z)$. Hence

$$\mathcal{L}(z) \subset \mathcal{V}_2(z).$$

Thus, by taking the convex hull of both sides, being $\mathcal{V}_2(z)$ convex, yields $\mathrm{co}\,\mathcal{L}(z) \subset \mathcal{V}_2(z)$, that is (2.11), and this finishes the proof. ■

Building on the conditions given by the above result and to the fact that, thanks to Theorem 2.1, the search of the set $\mathcal{A}$ can be carried out by focusing on a sublevel set of a certain quadratic function, the next result gives a first sufficient condition to solve Problem 2.1.

Proposition 2.1. *If there exist $P \in \mathcal{S}^n_+$, $S_1, S_2 \in \mathcal{D}^m_+$, and a positive scalar τ such that*

$$\mathcal{N} = \begin{bmatrix} \mathrm{He}(P(A + BK)) + \tau P & PB - K^\mathsf{T} S_2 \\ \bullet & -S_1 - 2S_2 \end{bmatrix} < \mathbf{0} \tag{2.14}$$

$$\mathrm{trace}(S_1)\Delta^2 - \tau \leq 0 \tag{2.15}$$

[1]Positive definiteness of S_1 implies that the function $v \mapsto v^\mathsf{T} S_1 v - \mathrm{trace}(S_1)\Delta^2$ is convex, then its sublevel sets are convex sets; see, *e.g.*, [14].

then,

$$\mathcal{A} = \mathcal{E}(P) \tag{2.16}$$

solves Problem 2.1.

Proof. For every $x \in \mathbb{R}^n$, consider the following quadratic function $V(x) = x^\mathsf{T} P x$. Following the ideas presented in the proof of Theorem 2.1, we want to prove that under (2.14) and (2.15) there exists a positive real scalar β such that

$$\langle \nabla V(x), w\rangle \leq -\beta V(x) \qquad \forall x \in \mathbb{R}^n \setminus \mathrm{Int}\mathcal{A}, w \in \mathcal{K}[X](x). \tag{2.17}$$

As the above relation is analogous to (2.8) in the proof of Theorem 2.1, establishing (2.17) suffices to show that the set $\mathcal{A}$ in (2.16) is UGAS for (2.4). By S-procedure arguments, (2.17) can be verified by showing that for every $x \in \mathbb{R}^n$, there exists a positive real scalar τ such that

$$\langle \nabla V(x), w\rangle - \tau(1 - x^\mathsf{T} P x) \leq -\beta V(x) \qquad \forall w \in \mathcal{K}[X](x). \tag{2.18}$$

On the other hand, via Proposition 1.1 and Proposition 1.2, for every $w \in \mathcal{K}[X](x)$, there exists $v \in \mathcal{K}[\Psi](Kx)$, such that $w = (A+BK)x + Bv$. Then, still by S-procedure arguments and according to Lemma 2.2, (2.18) is ensured by proving that for each $x \in \mathbb{R}^n$, and for each $v \in \mathbb{R}^m$,

$$\begin{aligned} &\langle \nabla V(x), (A+BK)x + Bv\rangle - \tau(1 - x^\mathsf{T} P x) - v^\mathsf{T} S_1 v \\ &+ \mathrm{trace}(S_1)\Delta^2 - 2v^\mathsf{T} S_2 (v + Kx) \leq -\beta V(x). \end{aligned} \tag{2.19}$$

By straightforward calculations the left-hand side of the above relation can be rewritten as follows

$$\begin{bmatrix} x \\ v \end{bmatrix}^T \mathcal{N} \begin{bmatrix} x \\ v \end{bmatrix} + \mathrm{trace}(S_1)\Delta^2 - \tau. \tag{2.20}$$

Thus in view of (2.14) and (2.15), it follows that there exists a small enough positive scalar γ such that for every $x \in \mathbb{R}^n \setminus \mathrm{Int}\mathcal{A}, w \in \mathcal{K}[X](x)$, one has $\langle \nabla V(x), w\rangle \leq -\gamma x^\mathsf{T} x$. Then, since for every $x \in \mathbb{R}^n$, $V(x) \leq \lambda_{\max}(P) x^\mathsf{T} x$, by setting $\beta = \frac{\gamma}{\lambda_{\max}(P)}$ gives (2.18), and this finishes the proof. ■

Remark 2.1. In the proof of the above result, we relied on Proposition 1.2 to build an overapproximation of $\mathcal{K}[X]$, avoiding the derivation of the exact expression of $\mathcal{K}[X]$, that is in general a nontrivial task. However, as argued in Remark 1.3, whenever $\mathrm{rank}\, K = m$ such an expression could be obtained by following similar arguments to [97, Theorem 1] and by relying on Proposition 1.3. On the one hand, due to the approach we embrace, following this approach would not give rise to any change in the derived conditions (the same sector conditions would be considered also in this case). On the other hand, the derivation of the actual Krasovskii regularization of $x \mapsto \Psi(Kx)$ could allow, in some case, a deep understanding of the dynamics of (2.4). This aspect will be clarified in Section 2.2.5 via some numerical examples.

The above result provides a sufficient condition to solve Problem 2.1. A necessary condition to ensure the feasibility of (2.14) is that the matrix $A+BK$ is Hurwitz. On the other hand, from Theorem 2.1, it turns out that having $A+BK$ Hurwitz enables to exhibit a solution to Problem 2.1. Therefore, at a first sight, the conditions provided by Proposition 2.1 could appear stronger than the mere Hurwitzness of the matrix $A+BK$. In other words, one may wonder whether the Hurwitzness of the matrix $A+BK$ ensures the feasibility of conditions (2.14) and (2.15). A positive answer is given by the following result.

Proposition 2.2. *Let $K \in \mathbb{R}^{m\times n}$ such that $A+BK$ is Hurwitz. Then, there exists $(\tau, P, S_1, S_2) \in \mathbb{R}_{>0} \times \mathcal{S}_+^n \times \mathcal{D}_+^n \times \mathcal{D}_+^n$ satisfying (2.14) and (2.15).*

Proof. Assume there exist $(\overline{\tau}, \overline{P}, \overline{S_1}) \in \mathbb{R}_{>0} \times \mathcal{S}_+^n \times \mathcal{D}_+^m$ such that

$$\begin{bmatrix} \mathrm{He}(\overline{P}(A+BK)) + \overline{\tau}\overline{P} & \overline{P}B \\ \bullet & -\overline{S}_1 \end{bmatrix} < \mathbf{0} \tag{2.21}$$

$$\mathrm{trace}(\overline{S_1})\Delta^2 - \overline{\tau} \leq 0. \tag{2.22}$$

For every diagonal $S_2 \in \mathbb{R}^{p\times p}$, define

$$\mathcal{M}(S_2) := \begin{bmatrix} \mathrm{He}(\overline{P}(A+BK)) + \overline{\tau}\overline{P} & \overline{P}B - K^\mathsf{T} S_2 \\ \bullet & -\overline{S}_1 - 2S_2 \end{bmatrix}.$$

From (2.21) it follows that $\mathcal{M}(\mathbf{0}) < \mathbf{0}$. Moreover, since $\mathcal{M}(S_2)$ depends continuously on the entries of S_2, there exists a small enough positive scalar δ, such that for every $S_2 \in \delta\mathcal{D}_+^m$ with $S_2 \leq \delta \mathrm{I}$ yields[2] $\mathcal{M}(S_2) < \mathbf{0}$.

To conclude the proof, it suffices to show that whenever $A+BK$ is Hurwitz there exists $(\overline{\tau}, \overline{P}, \overline{S}_1) \in \mathbb{R}_{>0} \times \mathcal{S}_+^n \times \mathcal{D}_+^m$ such that (2.21) and (2.22) holds. To this end, define $A_{cl} = A+BK$, and let $\mathcal{R}(A_{cl}) := \{|\Re(\lambda)|\colon \lambda \in \mathrm{spec}(A_{cl})\}$, notice that since A_{cl} is Hurwitz, then $\mathcal{R}(A_{cl}) \subset \mathbb{R}_{>0}$. Pick $\overline{\tau} \in (0, 2\min \mathcal{R}(A_{cl}))$, and define, $\widetilde{A}_{cl} = A_{cl} + \frac{\overline{\tau}}{2}\mathrm{I}$. Observe that, according to the selection considered for $\overline{\tau}$, $\widetilde{A}_{cl}$ is Hurwitz. Select $\overline{S}_1 \in \mathcal{D}_+^m$, such that $\mathrm{trace}(\overline{S}_1)\Delta^2 - \overline{\tau} \leq 0$. By following these choices, the right-hand side of (2.21) reads

$$\begin{bmatrix} \mathrm{He}(\widetilde{A}_{cl}^\mathsf{T} P) & PB \\ \bullet & -\overline{S}_1 \end{bmatrix}. \tag{2.23}$$

For any $\overline{Q}_1 \in \mathcal{S}_+^n$, pick the solution $\overline{W} \in \mathcal{S}_+^n$ to the following matrix equality

$$\mathrm{He}(\widetilde{A}_{cl}\overline{W}) = -B\overline{S}_1^{-1}B^\mathsf{T} - \overline{Q}_1$$

notice that such a solution always exists since $\widetilde{A}_{cl}$ is Hurwitz, and $\overline{S}_1 \in \mathcal{D}_+^m$. Now, set in

[2]This fact can be justified by noticing that the set $H := \{v \in \mathbb{R}^m \colon \mathcal{M}(\mathrm{diag}\{v_1, v_2, \ldots, v_m\}) < \mathbf{0}\}$ is open. Then, since $0 \in H$, there exists a positive scalar ε such that $\varepsilon\mathbb{B} \subset H$. Thus, by picking $\delta = \frac{1}{\sqrt{m}}\varepsilon$ yields the result.

(2.23), $P = \overline{W}^{-1}$. By following this choice, (2.23) becomes

$$\begin{bmatrix} \mathrm{He}(\widetilde{A}_{cl}^{\mathsf{T}}\overline{W}^{-1}) & \overline{W}^{-1}B \\ \bullet & -\overline{S}_1 \end{bmatrix}. \tag{2.24}$$

We want to show that the latter matrix is negative definite. By pre-and-post multiplying (2.24) by $\mathrm{diag}(\overline{W}, \mathrm{I})$, it turns out that (2.24) is negative definite if and only if

$$\begin{bmatrix} \mathrm{He}(\widetilde{A}_{cl}\overline{W}) & B \\ \bullet & -\overline{S}_1 \end{bmatrix} < \mathbf{0} \tag{2.25}$$

and the latter, due to the selection done for $\overline{W}$ turns into

$$\begin{bmatrix} -B\overline{S}_1^{-1}B^{\mathsf{T}} - \overline{Q}_1 & B \\ \bullet & -\overline{S}_1 \end{bmatrix} < \mathbf{0} \tag{2.26}$$

Moreover, by Schur complement, as $\overline{S}_1$ is positive definite, (2.26) is negative definite if and only if

$$-B\overline{S}_1^{-1}B^{\mathsf{T}} - \overline{Q}_1 + B\overline{S}_1^{-1}B = -\overline{Q}_1 < \mathbf{0} \tag{2.27}$$

which is obviously satisfied being $\overline{Q}_1 \in \mathcal{S}_+^n$. Then, $(\overline{\tau}, \overline{W}^{-1}, \overline{S}_1)$ establishes the result. ■

Remark 2.2. Notice that, whenever τ is fixed, (2.14) and (2.15) are linear in the decision variables. Therefore, Proposition 2.1 turns the solution to Problem 2.1 into a "quasi"-LMI feasibility problem. These aspects will be clarified in the sequel.

2.2.3 Controller Design

In the previous section of this chapter, we focused on the analysis problem of the quantized closed-loop system (2.4). Essentially, building on a stabilizing state-feedback controller for the quantization free closed-loop system, we shown that there exists a compact set $\mathcal{A}$ surrounding the origin which is UGAS for the closed-loop system. Such a set may contain limit-cycles and or parasitic equilibria for the closed-loop system that are undesired behaviors in engineered systems. Then, with the aim of limiting the influence of these phenomena, one may want to design the controller K so as to shrink the size of the set $\mathcal{A}$. To this end, in this section we propose certain constructive conditions characterizing the solutions to the problem formalized as follows.

Problem 2.2. (Controller design) Let A, B be matrices of adequate dimensions. Determine a gain $K \in \mathbb{R}^{m\times n}$, and a compact set $\mathcal{A} \subset \mathbb{R}^n$ containing the origin, such that $\mathcal{A}$ is UGAS for system (2.4).

At a first sight, Problem 2.2 could be solved directly by searching for a feasible solution to conditions (2.14) and (2.15), with the only caveat to treat also K as a variable. On the other hand, (2.14) and (2.15) are nonlinear in the decision variables. Hence, from a numerical standpoint, Proposition 2.1 does not provide an effective solution to Problem 2.2.

To overcome this problem, let us consider the following result.

Proposition 2.3. *If there exist $W \in \mathcal{S}_+^n$, $S_1, S_2 \in \mathcal{D}_+^m$, $Y \in \mathbb{R}^{m\times n}$, and a positive scalar τ, such that (2.15) is verified and,*

$$\begin{bmatrix} \mathrm{He}(AW+BY)+\tau W & B-Y^{\mathsf{T}}S_2 \\ \bullet & -S_1-2S_2 \end{bmatrix} < \mathbf{0} \tag{2.28}$$

then,

$$\mathcal{A} = \mathcal{E}(W^{-1}) \tag{2.29}$$

$$K = YW^{-1} \tag{2.30}$$

solve Problem 2.2.

Proof. The proof of this result is based on Proposition (2.1). In fact, we prove that condition (2.28) is obtained from (2.14) by means of a congruence transformation and an invertible change of variable. Let us assume that (2.28) is verified. Then, since from (2.30), $YW^{-1} = K$, pre-and-post multiplying the right-hand side of (2.28) by $\mathrm{diag}(W^{-1}, \mathrm{I})$, yields

$$\begin{bmatrix} \mathrm{He}(W^{-1}A+W^{-1}BK)+\tau W^{-1} & W^{-1}B-K^{\mathsf{T}}S_2 \\ \bullet & -S_1-2S_2 \end{bmatrix} < \mathbf{0}.$$

Finally, by setting in the previous relation $W^{-1} = P$ yields (2.14). Hence, thanks to Proposition 2.1 the assert is proven. ∎

Remark 2.3. Although the above result alleviates one of the nonlinearity affecting condition (2.14), (2.28) is still nonlinear in the decision variables. This aspect will be discussed in the sequel.

Clearly, as shown for Proposition 2.1, also in this case the feasibility of the conditions given by Proposition 2.5 is always ensured (under Assumption 1.2 on Page 21). In this sense, let us consider the following result that follows directly from Proposition 2.2.

Proposition 2.4. *Let A, B matrices such that Assumption 1.2 is satisfied. Then, there exists $(\tau, W, S_1, S_2, Y) \in \mathbb{R}_{>0} \times \mathcal{S}_+^n \times \mathcal{D}_+^n \times \mathcal{D}_+^n \times \mathbb{R}^{m\times n}$ satisfying (2.28) and (2.15).*

Proof. Since from Assumption 1.2 the pair A, B is stabilizable, there exists a gain K such that $A + BK$ is Hurwitz. Then, since condition (2.28) is obtained from condition (2.14) via invertible changes of variables and congruence transformations, by following the same steps as in the proof of Proposition 2.2, and by setting $Y = KW$ yields the result. ∎

2.2.4 Optimization Issues

It appears obvious that in solving Problem 2.1, one looks for an UGAS set which mostly fits the real behavior of the closed-loop system. On the other hand, in solving Problem 2.2, the main objective consists of designing the gain K to ensure that the closed-loop solutions stay sufficiently close to the origin. To this end, building on the conditions provided by Proposition 2.1 and by Proposition 2.3, one can consider the two following optimization problems:

Problem 2.3 (Stability)**.** Let A, B, K be matrices of adequate dimensions. Determine $P \in \mathcal{S}_+^n$, such that $\mathcal{E}(P)$ is UGAS for system (2.4), and it is minimized with respect to some criterion.

Problem 2.4 (Stabilization)**.** Let A, B be matrices of adequate dimensions. Determine a gain $K \in \mathbb{R}^{m\times n}$, and $P \in \mathcal{S}_+^n$, such that $\mathcal{E}(P)$ is UGAS for system (2.4), and it is minimized with respect to some criterion.

Notice that, although the two above problems are formulated in a similar fashion, they are in fact quite different. Indeed, in solving Problem 2.3, one attempts to reduce the conservatism in the analysis of the closed-loop system behavior. Instead, solving Problem 2.4 means to actively act on the closed-loop system by designing the controller gain K, to impose a desired behavior. The solution to the two above optimization problems can be carried out by embedding the conditions provided, respectively, by Proposition 2.1, and Proposition 2.3 into a suitable optimization scheme. To this end, an adequate measure of the sets $\mathcal{E}(P)$ and $\mathcal{E}(W^{-1})$ needs to be selected. Namely, the objective consists of defining a function $\mathcal{M}_a\colon \mathbb{R}^{n\times n} \to \mathbb{R}$ ($\mathcal{M}_s\colon \mathbb{R}^{n\times n} \to \mathbb{R}$), such that $\mathcal{M}_a(P)$ ($\mathcal{M}_s(W)$) provides a convenient indication on the size of $\mathcal{E}(P)$ ($\mathcal{E}(W^{-1})$). Once $\mathcal{M}_a$ ($\mathcal{M}_s$) is defined, Proposition 2.1 (Proposition 2.3) enables to reformulate Problem 2.3 (Problem 2.4) as follows:

$$\begin{aligned}
&\underset{P,S_1,S_2,\tau}{\text{minimize}} && \mathcal{M}_a(P)\\
&\text{subject to} && S_1, S_2 \in \mathcal{D}_+^m, P \in \mathcal{S}_+^n, \tau > 0\\
& && (2.14), (2.15).
\end{aligned} \tag{2.31}$$

$$\begin{aligned}
&\underset{W,S_1,S_2,\tau,Y}{\text{minimize}} && \mathcal{M}_s(W)\\
&\text{subject to} && S_1, S_2 \in \mathcal{D}_+^m, W \in \mathcal{S}_+^n, \tau > 0\\
& && (2.28), (2.15).
\end{aligned} \tag{2.32}$$

Size Criteria

Being the considered set, in both the above optimization problems, an ellipsoidal set, several criteria can be adopted to obtain a measure of such a set; see, *e.g.*, [15, 66, 120]. A first choice is to consider the volume of $\mathcal{E}(P)$ ($\mathcal{E}(W^{-1})$) as size criterion, *i.e.*, $\mathbf{vol}(\mathcal{E}(P))$ ($\mathbf{vol}(\mathcal{E}(W^{-1}))$). In particular, it turns out that, given $S \in \mathcal{S}_+^n$, and a generic ellipsoidal set $\mathcal{E}(S) := \{w \in \mathbb{R}^n\colon w^\mathsf{T} S w \leq 1\}$, then $\mathbf{vol}\,(\mathcal{E}(S)) \propto \sqrt{\det(S^{-1})}$; see [15]. Thus, adopting this criterion leads to $M_a(P) = -\det(P)$ and $M_s(W) = \det(W)$. However, as the two functions $M_a(P) = -\det(P)$ and $M_s(W) = \det(W)$ are in general non-convex, this would lead to possible N-P hard problems; see [15]. Therefore, with the aim of obtaining a numerically tractable optimization problem, the above criteria cannot be adapted directly.

Concerning Problem 2.3, a straightforward strategy to overcome this drawback, see [15]), consists of considering, as objective function to minimize $-\log\det(P)$. Indeed, the function $-\log\det(P)$ is convex over the set $\mathcal{S}_+^n$ and its minimization is equivalent to the minimization of $M_a(P) = -\det(P)$; see [15]. On the other hand, the adoption of the latter criterion could lead to a set $\mathcal{E}(P)$ excessively stretched along some direction. This is a well known behavior in the literature; see [120]. To overcome this problem, instead of minimizing the volume of $\mathcal{E}(P)$, one can minimize the trace(P^{-1}). Indeed, as trace$(P^{-1}) = \sum_{i=1}^n \lambda_i(P^{-1})$, $P > \mathbf{0}$, and each eigenvalue of P^{-1} corresponds to length of one of the axis of the ellipsoid $\mathcal{E}(P)$, minimizing trace(P^{-1}) tends to homogeneously shrink the set $\mathcal{E}(P)$ is each direction. However, since this criterion is in general non convex in the decision variable P, its exploitation in a numerical scheme is not straightforward. To overcome this drawback, we introduce a further variable $N \in \mathcal{S}_+^n$, subject to the following linear constraint

$$\begin{bmatrix} N & \mathrm{I} \\ \bullet & P \end{bmatrix} \geq \mathbf{0}$$

which, by Schur complement, is equivalent to $P^{-1} \leq N$. Therefore, the minimization of

$$\mathrm{trace}(P^{-1})$$

can be implicitly performed by minimizing trace(N), which is a convex (in fact linear) function of N. By pursuing this approach, Problem 2.3 reads

$$\begin{aligned} \underset{P,S_1,S_2,\tau,N}{\text{minimize}} \quad & \mathrm{trace}(N) \\ \text{subject to} \quad & \begin{bmatrix} N & \mathrm{I} \\ \bullet & P \end{bmatrix} \geq \mathbf{0} \\ & S_1, S_2 \in \mathcal{D}_+^m, P, N \in \mathcal{S}_+^n, \tau > 0 \\ & (2.14), (2.15). \end{aligned} \tag{2.33}$$

Another alternative solution, inspired from [120, 66], and that can be used to state Problem 2.3, consists of minimizing the set $\mathcal{E}(P)$ along certain directions of interests, (this method does not directly requires to specify a measure for the considered sets). In particular, let $v_1, v_2, \ldots, v_p \in \mathbb{R}^n$ be some given vectors, and let $\theta_1, \theta_2, \ldots, \theta_p$, positive scalars. Consider for each $i = 1, 2, \ldots, p$, the following constraints

$$v_i^\mathsf{T} P v_i \geq \theta_i \qquad i = 1, 2, \ldots, s. \tag{2.34}$$

By maximizing the scalars θ_i, the set $\mathcal{E}(P)$ shrinks along the directions v_i.

Hence, *e.g.*, via a linear scalarization, the above size criterion, can be adopted to state

Problem 2.3 as single objective optimization problem, as follows

$$\begin{aligned}
\underset{P,S_1,S_2,\tau,\theta_1,\theta_2,\ldots,\theta_s}{\text{minimize}} \quad & -\sum_{i=1}^{s}\theta_i\gamma_i \\
\text{subject to} \quad & S_1, S_2 \in \mathcal{D}_+^m, P \in \mathcal{S}_+^n, \tau > 0 \\
& (2.15), (2.14), (2.34)
\end{aligned} \tag{2.35}$$

where $\gamma_i > 0$ are the weights of the objectives.

Even in Problem 2.4, the above trace criterion can be easily adopted, and its exploitation is also simpler than in Problem 2.3; indeed it suffices to consider as convex objective in the decision variables directly $\operatorname{trace}(W)$. In particular, this choice leads to the following optimization problem

$$\begin{aligned}
\underset{W,S_1,S_2,\tau,Y}{\text{minimize}} \quad & \operatorname{trace}(W) \\
\text{subject to} \quad & S_1, S_2 \in \mathcal{D}_+^m, W \in \mathcal{S}_+^n, \tau > 0 \\
& (2.28), (2.15).
\end{aligned} \tag{2.36}$$

However, if one insists in requiring convexity for the measure criterion, adopting the above illustrated volume criterion is impossible. Indeed, the function $\log\det(W)$ is concave.

Numerical Issues in the Solution to (2.31)

Concerning (2.31), notice that, as long as the considered objective function is convex, whenever the scalar τ is fixed, such a problem is a genuine convex optimization problem over LMI constraints. Then, the solution to this problem can be performed in polynomial time via interior points methods; see [15]. On the other hand, the positive scalar τ can be treated as a tuning parameter, or being selected via an iterative search. This is a typical scenario in the literature; see, *e.g.*, [118, 119, 126]. Then (2.31) can be efficiently solved on a computer, with only caveat to obtain a sub-optimal solution. Based on this idea, consider the following algorithm that, by performing a grid search for τ in an interval wherein the feasibility of (2.14)-(2.15) is ensured, provides a possible solution to (2.31)

Algorithm 2.1 Stability analysis

Input: Matrices A, B, K, scalars $\Delta > 0$, a convex function $\mathcal{M}_a \colon \mathcal{S}_+^n \to \mathbb{R}_{>0}$, and a tolerance $\rho > 0$.

Initialization: Let $\mathcal{R}(A + BK) := \{|\Re(\lambda)| \colon \lambda \in \operatorname{spec}(A + BK)\}$, select $\tau = 2 \times 0.99 \min \mathcal{R}(A + BK)$,

Iteration

Step 1:

Solve the following convex optimization problem over LMIs

$$
\begin{aligned}
&\underset{S_1, S_2, P}{\text{minimize}} && \mathcal{M}_a(P) \\
&\text{s.t.} && S_1, S_2 \in \mathcal{D}_+^m, P \in \mathcal{S}_+^n \\
& && \begin{bmatrix} \operatorname{He}(P(A + BK)) + \tau P & PB - K^{\mathsf{T}} S_2 \\ \bullet & -S_1 - 2S_2 \end{bmatrix} < \mathbf{0} \\
& && \operatorname{trace}(S_1)\Delta^2 - \tau \leq 0
\end{aligned}
$$

Pick the sub-optimal solution $(\overline{P}, \overline{S}_1, \overline{S}_2)$. Store the obtained solution:

$$
\mathcal{M}_{a\star}^{(k)} \leftarrow \mathcal{M}_a(\overline{P}), P_\star^{(k)} \leftarrow \overline{P}.
$$

$k \leftarrow k + 1$

Step 2:

Decrease τ of ρ, *i.e.*, $\tau \leftarrow \tau - \rho$

Until $\tau > 0$.

Step 3: $k_{\max} \leftarrow k$, select $k^\star = \underset{k \in \{1, 2, k_{\max}\}}{\operatorname{argmin}} \{\mathcal{M}_{a\star}^{(k)}\}$

Output: $P = P_\star^{(k^\star)}$.

Remark 2.4. Notice that, as shown in the proof of Proposition 2.2, the initialization proposed ensures that at each iteration, Step 1 terminates with a sub-optimal solution to Problem 2.3. Then, Algorithm 2.1 always terminates in a finite number of steps.

Numerical Issues in the Solution to (2.32)

The solution to (2.32) is much more complicated, and for this further work is needed. Indeed, even when τ is fixed, condition (2.28) is nonlinear due to the product of decision variables $Y^{\mathsf{T}} S_2$ (and its transpose). This kind of nonlinearity often occurs whenever one attempts to design, via the solution of an optimization problem, a static state feedback controller for certain class of nonlinear systems; see, *e.g.*, [118]. Nevertheless, being S_2 diagonal, at least for $m \leq 2$, even for this variable a grid search can be envisaged to solve (2.32), with still the only caveat to obtain a suboptimal solution.

Another strategy to ride over this problem consists of adopting a procedure indicated here below:

- As a first step one selects some stabilizing gain for the pair (A, B), this is always possible due to Assumption 1.2
- Once the controller gain is known, by fixing τ as prescribed in the proof of Proposition 2.2, (2.28) becomes a genuine LMI in the remaining variables, whose feasible set is non-empty. Therefore, S_2 can be selected to ensure the feasibility of (2.28)-(2.15)
- Once S_2 is selected as indicated above, by preforming a grid search for τ, a suboptimal solution to (2.32) can be determined by solving a finite number of convex optimization problems over LMIs.

These steps are exploited to build the following algorithm.

Algorithm 2.2 Controller design

Input: Matrices A, B, scalar $\Delta > 0$, a tolerance $\rho > 0$, and a convex function $\mathcal{M}_s \colon \mathcal{S}_+^n \to \mathbb{R}_{>0}$.

Initialization: Select K, such that $A+BK$ is Hurwitz. Let $\mathcal{R}(A+BK) \coloneqq \{|\Re(\lambda)| \colon \lambda \in \operatorname{spec}(A+BK)\}$. Set for the next step

$$\overline{\tau} = 2 \times 0.99 \min \mathcal{R}(A+BK)$$

Step 1:

Determine a feasible solution to the following LMI problem

$$\begin{aligned} & S_1, S_2 \in \mathcal{D}_+^m, P \in \mathcal{S}_+^n \\ & \begin{bmatrix} \operatorname{He}(P(A+BK)) + \overline{\tau} P & PB - K^\mathsf{T} S_2 \\ \bullet & -2S_2 - S_1 \end{bmatrix} < \mathbf{0} \\ & \operatorname{trace}(S_1)\Delta^2 - \overline{\tau} \le 0 \end{aligned}$$

Set $\overline{S}_2 = S_2$ for the next step. Select a grid of positive values $\mathcal{G}_\tau$ such that $\overline{\tau} = \max \mathcal{G}_\tau$

Iteration

Step 2:

Solve the following convex optimization problem over LMIs selecting τ over $\mathcal{G}_\tau$

$$\begin{aligned} \underset{W, S_1, Y}{\text{minimize}} \quad & \mathcal{M}_s(W) \\ \text{subject to} \quad & S_1 \in \mathcal{D}_+^m, P \in \mathcal{S}_+^n \\ & \begin{bmatrix} \operatorname{He}(AW + BY) + \overline{\tau} W & B - Y^\mathsf{T} \overline{S}_2 \\ \bullet & -2\overline{S}_2 - S_1 \end{bmatrix} < \mathbf{0} \\ & \operatorname{trace}(S_1)\Delta^2 - \overline{\tau} \le 0 \end{aligned}$$

Pick the suboptimal solution to the above optimization problem

$$(\tau^\star, W^\star, Y^\star, S_1^\star).$$

and determine the controller gain as $K^\star = Y^\star (W^\star)^{-1}$.

Determine the closed-loop matrix $A+BK^\star$, and set $\overline{\tau} = 2 \times 0.99 \min \mathcal{R}(A+BK^\star)$. Build a grid of positive values $\mathcal{G}_\tau$ such that $\overline{\tau} = \max \mathcal{G}_\tau$, and $\tau^\star \in \mathcal{G}_\tau$, (notice that necessarily $\tau^\star \le \overline{\tau}$. Including $\tau^\star$ in $\mathcal{G}_\tau$ ensures the feasibility at the next step).

Until $\mathcal{M}_s(W)$ does not decrease below ρ over three consecutive steps.

Output: $(K^\star, P = (W^\star)^{-1})$

Remark 2.5. The above algorithm essentially performs a grid search for τ keeping the value of S_2 unchanged from the initialization stage. However, the grid search proposed is "greedier" than a standard one. Indeed, the grid $\mathcal{G}$ is built from scratch at each iteration, to tentatively explore a wider portion of the feasible set, at least in the τ-direction.

Remark 2.6. The proposed algorithm has two important properties. The first one is that, thanks to the initialization proposed building on the proof of Proposition 2.2, the algorithm always provides a suboptimal solution to the controller design problem. The second one is that, since at each iteration the objective is at least non-increasing, the algorithm stops in a finite number of iterations.

An alternative strategy to solve the controller design problem consists of exploiting the following sufficient condition to (2.28).

Proposition 2.5. *If there exist $W \in \mathcal{S}_+^n$, $S_1, H \in \mathcal{D}_+^m$, $Y \in \mathbb{R}^{m\times n}$, and a positive scalar τ, such that*

$$\begin{bmatrix} \mathrm{He}(AW+BY)+\tau W & BH-Y^\mathsf{T} & \mathbf{0} \\ \bullet & -4H & \mathrm{I} \\ \bullet & \bullet & -S_1 \end{bmatrix} < \mathbf{0} \tag{2.37}$$

then $W, \tau, Y, S_1, S_2 = H^{-1}$ satisfies (2.28).

Proof. By Schur complement, (2.37) implies

$$\begin{bmatrix} \mathrm{He}(AW+BY)+\tau W & BH-Y^\mathsf{T} \\ \bullet & -4H+S_1^{-1} \end{bmatrix} < \mathbf{0}. \tag{2.38}$$

On the other hand, being S_1 and H positive definite, one has

$$(H - S_1^{-1})S_1(H - S_1^{-1}) \geq \mathbf{0}$$

or equivalently

$$-2H + S_1^{-1} \geq -S_1H^2.$$

Then, it follows

$$\begin{bmatrix} \mathrm{He}(AW+BY)+\tau W & BH-Y^\mathsf{T} \\ \bullet & -2H-S_1H^2 \end{bmatrix} \leq \begin{bmatrix} \mathrm{He}(AW+BY)+\tau W & BH-Y^\mathsf{T} \\ \bullet & -4H+S_1^{-1} \end{bmatrix} < \mathbf{0}. \tag{2.39}$$

Moreover, pre-and-post multiplying the left-hand side of the above relation by $\mathrm{diag}(\mathrm{I}, H^{-1})$ yields

$$\begin{bmatrix} \mathrm{He}(AW+BY)+\tau W & B-Y^\mathsf{T}H^{-1} \\ \bullet & -2H^{-1}-S_1 \end{bmatrix} < \mathbf{0}. \tag{2.40}$$

Then, since setting $S_2 = H^{-1}$ yields (2.28), the assert is proven. ■

Thus, exploiting the above result, performing a grid search for the matrix S_2 (at least for $m \leq 2$), or a two-stage procedure, represent viable solutions to solve Problem 2.4 via a convex setup. Nevertheless, while the feasibility of the conditions provided by Proposition

2.3 is ensured by Proposition 2.4, there is no guarantees that Proposition 2.5 provides feasible conditions. Therefore, establishing which of the two techniques is more convenient is an open question.

A less evident aspect to be considered in solving (2.32) consists of avoiding solutions characterized by an overly large controller gain, situation that needs to be ruled out to envision the physical realization of the proposed controller. In particular, observe that the optimal solutions to (2.32) could in some case be approached only via an infinitely large controller gain. This phenomenon is thoroughly addressed in [110] for the case of static state-feedback $\mathcal{H}_\infty$-problem for linear systems. To overcome this problem, a typical solution consists of adding further constraints in (2.32) to limit the controller gain. This procedure somehow corresponds to reshape the feasible set of the considered optimization problem in way such that high-gain control solutions become unfeasible solutions. However, it follows from Proposition 2.3 that the matrix W is linked to both the set $\mathcal{E}(W^{-1})$ and to the gain K. Then, limiting the norm of the gain K by directly operating on the expression given in Proposition 2.3 leads to add further constrains on the matrix W. This fact may have a negative effect on the solution to (2.32). On the one hand, further constraining the matrix W may introduce an additional conservatism in the solution of (2.32). On the other hand, although the feasibility of (2.32) should not be affected by additional constraints on the matrix W, at least when those are not excessively severe, including further constraints on the matrix W may impact on the achievable suboptimal solutions. Loosely speaking, the addition of further constraints in the optimization problem can reshape the feasible set of (2.32) in a unfavorable fashion. To alleviate these issues, following the lines of [20], we provide a sufficient condition to (2.28) in which the matrices W and K are not directly coupled. In particular, let us consider the result given next

Corollary 2.1. If there exist $J \in \mathcal{S}_+^n$, $Y \in \mathbb{R}^{m\times n}$, $F \in \mathbb{R}^{n\times n}$, $S_1, S_2 \in \mathcal{D}_+^m$, and a positive scalar τ such that (2.15) is verified, and

$$\begin{bmatrix} -\operatorname{He}(F) & J + AF + BY - F^\mathsf{T} & B \\ \bullet & \tau J + \operatorname{He}(AF + BY) & -Y^\mathsf{T} S_2 + B \\ \bullet & \bullet & -S_1 - 2S_2 \end{bmatrix} < \mathbf{0} \tag{2.41}$$

then $K = YF^{-1}$ and $\mathcal{A} = \mathcal{E}(F^{-\mathsf{T}} J F^{-1})$ are solution to Problem 2.2.

Proof. The proof is inspired by [99]. From Proposition 2.1, notice that $\mathcal{N} = \mathcal{W}^\mathsf{T}\mathcal{Q}\mathcal{W}$, where

$$\mathcal{W} = \begin{bmatrix} A + BK & B \\ \mathrm{I} & \mathbf{0} \\ \mathbf{0} & \mathrm{I} \end{bmatrix}, \mathcal{Q} = \begin{bmatrix} \mathbf{0} & P & \mathbf{0} \\ \bullet & \tau P & -K^\mathsf{T} S_2 \\ \bullet & \bullet & -S_1 - 2S_2 \end{bmatrix}.$$

Thus, (2.14) can be rewritten equivalently as $\mathcal{W}^\mathsf{T}\mathcal{Q}\mathcal{W} < \mathbf{0}$. Moreover, being S_1 and S_2 positive definite, $\mathcal{U}^\mathsf{T}\mathcal{Q}\mathcal{U} < \mathbf{0}$, with $\mathcal{U}^\mathsf{T} = \begin{bmatrix}\mathbf{0} & \mathbf{0} & \mathrm{I}\end{bmatrix}$, is obviously satisfied. Thus, by the projection lemma; see [99], the satisfaction of (2.14), whenever S_1 and S_2 are required to be positive definite, is equivalent to find a matrix $\mathcal{X}$ such that

$$\mathcal{Q} + \mathcal{W}^{\mathsf{T}\perp}_{r}\mathcal{X}\mathcal{U}^{\perp}_{r} + \mathcal{U}^{\mathsf{T}\perp}_{r}\mathcal{X}^{\mathsf{T}}\mathcal{W}^{\perp}_{r} < \mathbf{0} \tag{2.42}$$

where, $\mathcal{U}^{\perp}_{r}$ and $\mathcal{W}^{\perp}_{r}$ are some matrices having as rows a basis of the row-null space, respectively of $\mathcal{U}$ and $\mathcal{W}$. Now, by selecting $\mathcal{U}^{\perp}_{r} = \begin{bmatrix} \mathrm{I}_{2n} & \mathbf{0}_{2n\times p} \end{bmatrix}$ and $\mathcal{W}^{\perp}_{r} = \begin{bmatrix} -\mathrm{I} & A+BK & B \end{bmatrix}$, and by partitioning $\mathcal{X} = \begin{bmatrix} X_1 & X_2 \end{bmatrix}$, where $X_1, X_2 \in \mathbb{R}^{n\times n}$, from (2.42) one gets

$$\begin{bmatrix} -\mathrm{He}(X_1) & P - X_2 + X_1^{\mathsf{T}}(A+BK) & X_1^{\mathsf{T}}B \\ \bullet & \mathrm{He}(X_2^{\mathsf{T}}(A+BK)) + \tau P & X_2^{\mathsf{T}}B - K^{\mathsf{T}}S_2 \\ \bullet & \bullet & -S_1 - 2S_2 \end{bmatrix} < \mathbf{0}. \tag{2.43}$$

At this stage, by setting in the above expression $X_1 = X_2 = X$, then by pre-and-post multiplying the left-hand side of the resulting matrix by $\mathrm{diag}(X^{-\mathsf{T}}, X^{-\mathsf{T}}, \mathrm{I})$ and $\mathrm{diag}(X^{-1}, X^{-1}, \mathrm{I})$ and finally by setting $X^{-1} = F$, $J = F^{\mathsf{T}}PF$ and $Y = KF$ yields the left-hand side of (2.41). Then, the satisfaction of (2.41) implies the satisfaction of (2.14). Therefore, thanks to Proposition 2.3, the assertion is proven. ∎

Remark 2.7. Notice that, the fact of choosing $X_1 = X_2$ in the derivation of the previous result adds some conservatism to the conditions given in Proposition 2.3. Specifically, differently from Proposition 2.3, there is no guarantees that the conditions provided by Proposition 2.7 are feasible.

Building from the previous result, with the objective of limiting the norm of the controller gain K, consider the result given next

Proposition 2.6. *If there exist two matrices $F \in \mathbb{R}^{n\times n}$, and $Y \in \mathbb{R}^{m\times n}$, and a positive scalar μ, such that*

$$\begin{bmatrix} \mathrm{He}(F) - \mathrm{I} & Y^{\mathsf{T}} \\ \bullet & \mu^2\mathrm{I} \end{bmatrix} \geq \mathbf{0} \tag{2.44}$$

then $\|YF^{-1}\| \leq \mu$.

Proof. First, from [33], $\mathrm{He}(F) - \mathrm{I} \leq F^{\mathsf{T}}F$, then (2.44) gives

$$\begin{bmatrix} F^{\mathsf{T}}F & Y^{\mathsf{T}} \\ \bullet & \mu^2\mathrm{I} \end{bmatrix} \geq \mathbf{0}. \tag{2.45}$$

Then, by pre-and-post multiplying the left-hand side of (2.45), respectively by, $\mathrm{diag}(F^{-\mathsf{T}}, \mathrm{I})$ and $\mathrm{diag}(F^{-1}, \mathrm{I})$, one gets

$$\begin{bmatrix} \mathrm{I} & F^{-\mathsf{T}}Y^{\mathsf{T}} \\ \bullet & \mu^2\mathrm{I} \end{bmatrix} \geq \mathbf{0}. \tag{2.46}$$

Then, by Schur complement (2.46) yields $F^{-\mathsf{T}}Y^{\mathsf{T}}YF^{-1} \leq \mu^2\mathrm{I}$, which in turn is equivalent to $\|YF^{-1}\| \leq \mu$, concluding the proof. ∎

Another strategy to, implicitly, limit the norm of the gain K consists of constraining the eigenvalues of the matrix $(A+BK)$ to lay in a suitable region contained in the open left-half complex plane. This kind of additional constraints can be easily expressed in a linear matrix inequality form; see [25]. A typical choice is to consider as region the closed circle centered in $(-\omega, 0)$ with radius $r > 0$, where ω is a positive real scalar, *i.e.*, $\{z \in \mathbb{C}\colon |z + \omega| \leq r\}$. Such a condition is guaranteed by considering the following constraint; see [118],

$$\begin{bmatrix} -rQ & \omega F + AF + BY \\ \bullet & -r(F^{\mathsf{T}} + F - Q) \end{bmatrix} < \mathbf{0} \tag{2.47}$$

where Q is a symmetric matrix with adequate dimensions.

As mentioned earlier, Corollary 2.1 allows to solve Problem 2.2, by decoupling the matrix defining the set $\mathcal{A} = \mathcal{E}(F^{-\mathsf{T}} J F^{-1})$, from the controller gain K. On the other hand, by doing so, the matrix defining the set $\mathcal{A}$ does not explicitly appear in (2.41). Then, embedding (2.41) into an optimization scheme to shrink the size of the set $\mathcal{A}$ requires further work. Suppose that one wants to insist on considering a trace criterion, that is minimizing $\mathrm{trace}(F J^{-1} F^{\mathsf{T}})$. A strategy that can be adopted to obtain a convex objective function to minimize consists of considering the following constraint

$$\begin{bmatrix} N & F \\ \bullet & J \end{bmatrix} \geq \mathbf{0} \tag{2.48}$$

where $N \in \mathcal{S}_+^n$. Indeed, the latter constraint is equivalent to $F J^{-1} F^{\mathsf{T}} \leq N$. Then, the minimization of $\mathrm{trace}(F J^{-1} F^{\mathsf{T}})$ can be performed indirectly via the minimization of $\mathrm{trace}(N)$. Therefore, Problem 2.4 can be formalized as follows

$$\begin{aligned} &\underset{F,J,S_1,S_2,Y,N}{\text{minimize}} && \mathrm{trace}(N) \\ &\text{s.t.} && S_1, S_2 \in \mathcal{D}_+^m,\ J, N \in \mathcal{S}_+^n \\ & && (2.41), (2.15), (2.48), (2.44)\ (\text{or } (2.47)), (2.48) \end{aligned} \tag{2.49}$$

Remark 2.8. Obviously, with the aim of limiting the norm of the controller gain, similar techniques as those illustrated above can be developed directly building from the conditions given by Proposition 2.3 (without the introduction of any slack variables) by adding further constraints on the matrix W. However, although the feasibility of the conditions given by Proposition 2.3 is ensured, whenever such conditions are coupled with further constraints, the feasibility of the resulting optimization problem cannot be ensured *a priori*. On the other hand, as mentioned earlier, due to the conservatism introduced by Corollary 2.1, even (2.49) could be unfeasible. Therefore, determining *a priori* which approach is the more convenient is an open question.

Similarly to Proposition 2.3, condition (2.41) is nonlinear in the decision variables. As matter of fact, condition (2.41) is affected by the same kind nonlinearities of condition (2.28), then the same techniques illustrated above can be used to alleviate these nonlinearities. In this sense, the result given next parallels Proposition 2.5.

Proposition 2.7. *If there exist* $W \in \mathcal{S}_+^n$*,* $S_1, H \in \mathcal{D}_+^m$*,* $Y \in \mathbb{R}^{m \times n}$*, and a positive scalar* τ*,* *such that*

$$\begin{bmatrix} -\operatorname{He}(F) & J + AF + BY - F^{\mathsf{T}} & B & \mathbf{0} \\ \bullet & \tau J + \operatorname{He}(AF + BY) & -Y^{\mathsf{T}} + BH & \mathbf{0} \\ \bullet & \bullet & -4H & \mathrm{I} \\ \bullet & \bullet & \bullet & -S_1 \end{bmatrix} < \mathbf{0} \tag{2.50}$$

then $J, F, \tau, Y, S_1, S_2 = H^{-1}$ *satisfies* (2.41).

Proof. The proof follows the same steps traced in the proof of Proposition 2.5, then it is omitted. ■

Remark 2.9. Notice that, although the solution to Problem 2.4 provides in one shot a solution to Problem 2.2, hence the controller gain, and the set $\mathcal{A}$, due to the further constraints introduced to render Problem 2.4 numerically tractable, the set $\mathcal{A}$ issued by this stage can be further tightened to fit more the behavior of the closed-loop system. Indeed, once the controller gain K is known, by performing an analysis stage via Proposition 2.1, further improvements can be obtained in terms of reduction of the size of set $\mathcal{A}$.

In the next section, the effectiveness of the proposed methodology is shown in some examples.

2.2.5 Numerical Examples

Example 2.1 (Furuta pendulum)**.** Consider the Furuta pendulum [67], whose linearized model around the unstable equilibrium point is given by

$$\dot{x} = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 39.32 & -14.52 & 0 \\ 0 & 81.78 & -13.98 & 0 \end{bmatrix} x + \begin{bmatrix} 0 \\ 0 \\ 25.54 \\ 24.59 \end{bmatrix} u \tag{2.51}$$

where x_1, x_2 represent respectively the base angle and the pendulum angle (rad), x_3 and x_4 are respectively the two angular speeds ($\mathrm{rad\,s^{-1}}$), and u is input voltage (V) of the motor driving the base shaft. Assume that the system is controlled via a static state feedback controller, with

$$K = \begin{bmatrix} 2.2710 & -27.1793 & 2.4963 & -3.9153 \end{bmatrix}$$

and that the actuator is quantized via uniform quantizer with $\Delta = 0.5$. By selecting as convex criterion $\mathcal{M}_a(P) = -\log\det(P)$, the solution of Problem 2.3, via the adoption of Algorithm 2.1, with a tolerance $\rho = 0.1$, yields

$$P = \begin{bmatrix} 20.6128 & -59.4021 & 7.79714 & -9.06105 \\ -59.4021 & 556.424 & -33.2957 & 40.6492 \\ 7.79714 & -33.2957 & 6.45171 & -7.62996 \\ -9.06105 & 40.6492 & -7.62996 & 10.3472 \end{bmatrix}$$

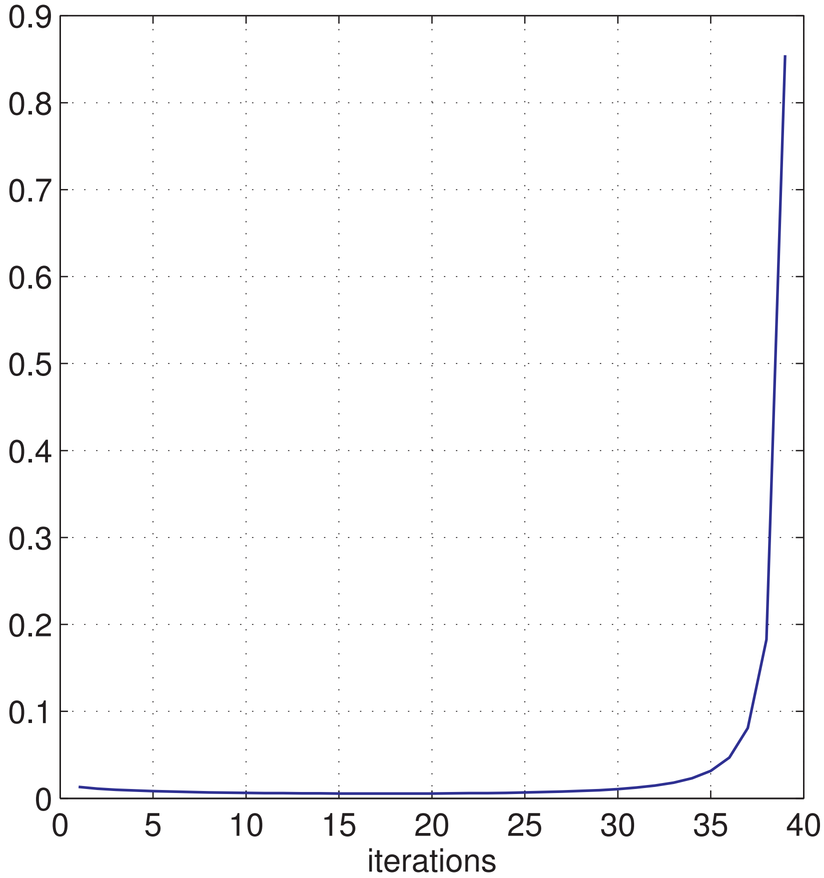


Figure 2.2: $\sqrt{\det(P^{-1})}$ versus the number of iterations.

Figure 2.2 shows the evolution of $\sqrt{\det(P)}$ (proportional to $\mathbf{vol}(\mathcal{E}(P))$) at each iteration of Algorithm 2.1.

To give a measure of the tightness of the set $\mathcal{A}$ with respect to the actual behavior of the closed-loop system, in Figure 2.3 we report the time-evolution of the function $x^{\mathsf{T}}Px$ along some solutions to the closed-loop system. The figure reveals that the trajectories once enter the set $\mathcal{A}$ (finite time convergence) no longer leave it and actually stay sufficiently close to its boundary.

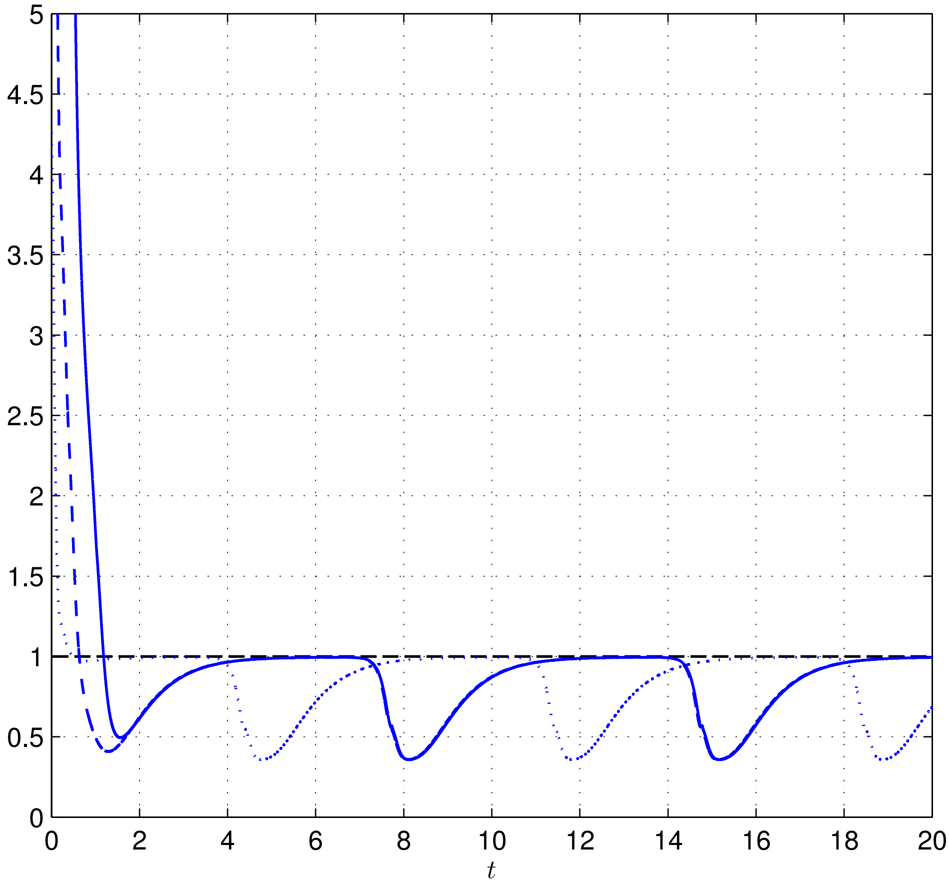


Figure 2.3: The evolution of the function $V(x) = x^{\mathsf{T}}Px$. $x_0 = (0, \pi/8, 0, 0)$ (solid-line), $x_0 = (0, \pi/18, 0, 0)$ (dashed-line), $x_0 = (0, \pi/36, 0, 0)$ (dotted-line).

Example 2.2. This example has the aim to show how the use of slack variables, as suggested in Corollary 2.1, can, in some cases, provide notable benefits. Consider the following data, borrowed from the balancing pointer system in [69], defining the closed-loop system in (2.4).

$$A = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, B = \begin{bmatrix} 0 \\ -1 \end{bmatrix}, \Delta = 1.$$

Assume that, we want to design a static state feedback controller by solving (2.36). Moreover, to avoid overly large controller gain, we limit the controller gain explicitly via a constraint like (2.44). In particular, by following the same steps in the proof of Proposition (2.6), it turns out that given $\mu > 0$, $\|K\| \leq \mu^2$ if

$$\begin{bmatrix} 2W - \mathrm{I} & Y^\mathsf{T} \\ \bullet & \mu\mathrm{I} \end{bmatrix} \geq \mathbf{0}$$

Therefore, pursuing this approach (2.36) reads

$$\begin{aligned} \underset{W,S_1,S_2,Y,R}{\text{minimize}} \quad & \operatorname{trace}(W) \\ \text{s.t.} \quad & S_1, S_2 \in \mathcal{D}_+^m, W \in \mathcal{S}_+^n \\ & (2.28), (2.15) \\ & \begin{bmatrix} 2W - \mathrm{I} & Y^\mathsf{T} \\ \bullet & \mu\mathrm{I} \end{bmatrix} \geq \mathbf{0}. \end{aligned} \tag{2.52}$$

For $\mu = 50, \tau = 0.99, S_2 = 0.1$ the solution of (2.52) yields, $K = \begin{bmatrix} 2.93 & 1.59 \end{bmatrix}$ and $\mathcal{A} = \mathcal{E}(W^{-1})$, with

$$W = \begin{bmatrix} 0.8009 & -0.4087 \\ -0.4087 & 1.151 \end{bmatrix}$$

for which one has $\operatorname{trace}(W) \approx 1.9521$.

Instead, solving (2.49), endowed with the additional constraint given by Proposition (2.6), still for $\mu = 50, \tau = 0.99, S_2 = 0.1$ provides

$$\begin{aligned} K &= \begin{bmatrix} 4.772 & 4.563 \end{bmatrix} \\ \mathcal{A} &= \mathcal{E}(F^{-\mathsf{T}} J F^{-1}) \end{aligned}$$

where

$$F^{-\mathsf{T}} J F^{-1} = \begin{bmatrix} 3.872 & 2.45 \\ 2.45 & 3.84 \end{bmatrix}$$

and for which $\operatorname{trace}(F J^{-1} F) \approx 0.869$. Namely the introduction of slack variables leads to an improvement of about 55.45% in terms of minimization of the size of $\mathcal{A}$, at least for the considered trace criterion. Figure 2.4 shows the two sets obtained by solving (2.49) and (2.52). The figures points out that, in this case, solving (2.49) enables to shrink more the size of the set $\mathcal{A}$.

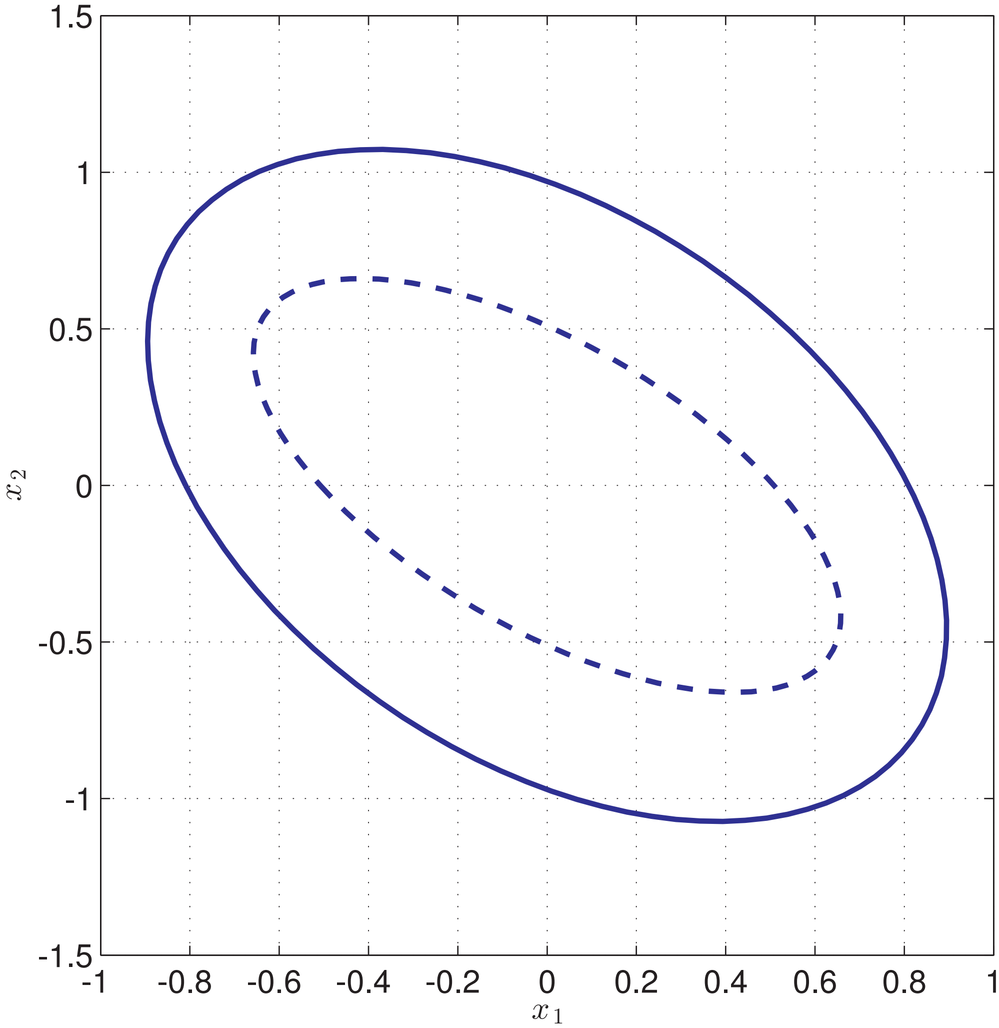


Figure 2.4: The two sets $\mathcal{A}$ resulting from the solution to the controller design problem. $\mathcal{E}(W^{-1})$ solid, $\mathcal{E}(F^{-\mathsf{T}}JF^{-1})$ dashed.

Example 2.3 (A multi-input system)**.** For the closed-loop system (2.4), consider the following example derived from [2] for which

$$A = \begin{bmatrix} -0.5 & 1.5 & 4 \\ 4.3 & 6 & 5 \\ 3.2 & 6.8 & 7.2 \end{bmatrix}, B = \begin{bmatrix} -0.7 & -1.3 \\ 0 & -4.3 \\ 0.8 & -1.5 \end{bmatrix}$$

and assume $\Delta = 0.5$. We solve (2.49) augmented with (2.47) for which $\omega = 10, r = 8.5$. To deal with the nonlinearities affecting (2.49), we select τ and S_2 over a three dimensional grid. In particular, the most convenient values selected are $\tau = 1.8, S_2 = \mathrm{diag}(1.4 \times 10^{-6}, 4.3 \times 10^{-5})$, giving

$$K = \begin{bmatrix} -0.71 & 1.9 & -27 \\ 4.3 & 4.1 & 4.3 \end{bmatrix}.$$

As a second step, to tighten more the set $\mathcal{A}$ obtained by the solution to (2.49), we perform an analysis stage via Algorithm 2.1, while considering as a convex criterion $\mathcal{M}_a(P) = -\operatorname{logdet}(P)$. Specifically, Algorithm 2.1 provides

$$P = \begin{bmatrix} 29.33 & 13.25 & -14.07 \\ 13.25 & 65 & -136.7 \\ -14.07 & -136.7 & 404.8 \end{bmatrix}.$$

Figure 2.5 shows the evolution of the closed-loop system in its state space, from different initial conditions. Furthermore, Figure 2.6 reports a particular closed-loop trajectory in its time-domain.

Simulations show that trajectories converge into the set $\mathcal{A} = \mathcal{E}(P)$. More precisely, closed-loop solutions appear to converge towards two equilibria contained in $\mathcal{E}(P)$. Notice that also the origin is an equilibrium point for the closed-loop system, though unstable. It is interesting to notice that these two equilibrium points appear to belong respectively to the two surfaces $Kx = [\Delta \ -\Delta]^\mathsf{T}$ and $Kx = [-\Delta \ \Delta]^\mathsf{T}$ wherein the function $\mathrm{q}(Kx)$ is discontinuous. As matter of fact, the two mentioned equilibria are actually Krasovskii equilibria, indeed there does not exist any point $\bar{x} \in \mathbb{R}^3$, such that

$$\begin{cases} A\bar{x} = -B \begin{bmatrix} \Delta \\ -\Delta \end{bmatrix} \\ K\bar{x} = \begin{bmatrix} \Delta \\ -\Delta \end{bmatrix} \end{cases}.$$

The same considerations hold for the other equilibrium point, for which $Kx = [-\Delta \ \Delta]^\mathsf{T}$. The determination of this kind of equilibria, in general, is far from trivial. However, in this example, the results provided by the above simulation may be used as a first guess to exactly determine the two Krasovskii equilibria. Specifically, let $\bar{x}$ be a Krasovskii equilibrium for the

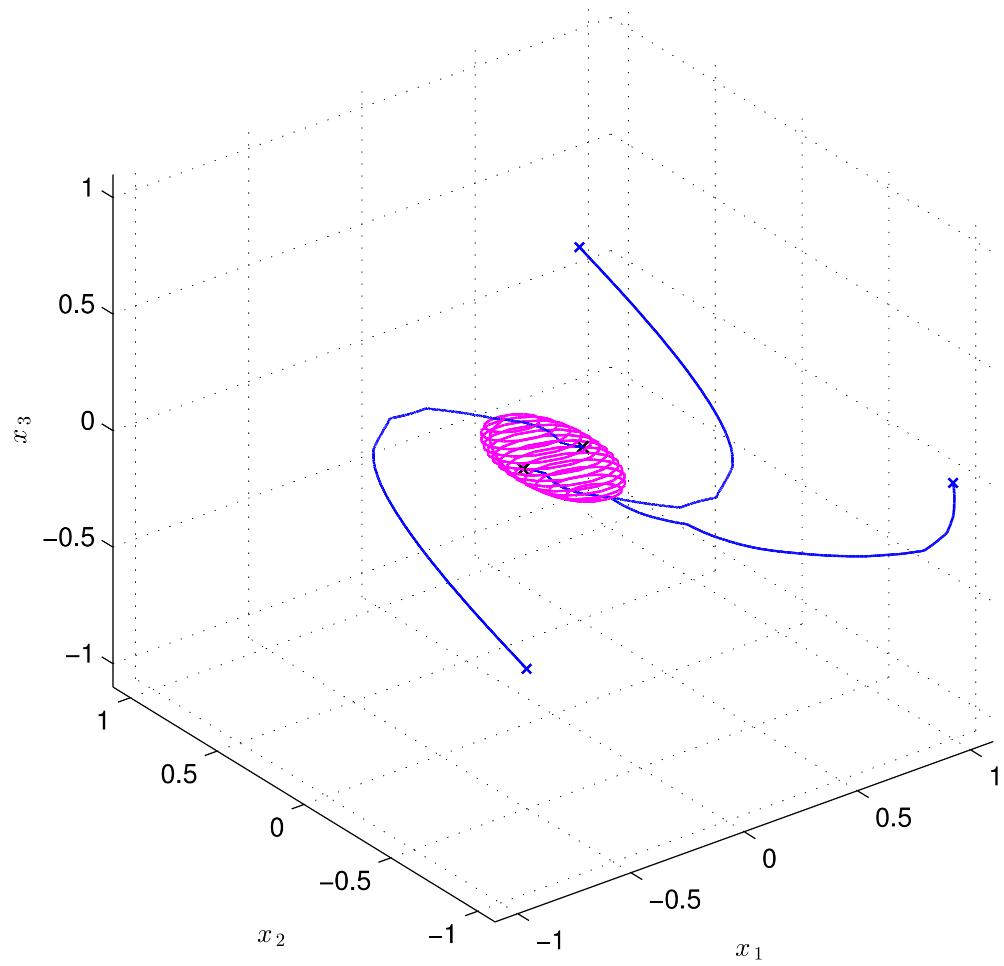


Figure 2.5: Some closed-loop solutions converging into the set $\mathcal{E}(P)$ (magenta). The solutions are obtained by integrating the closed-loop model via an Euler method with time step 10^{-4}.

closed-loop system and such that $K\bar{x} = [\Delta \ -\Delta]^{\mathsf{T}}$. By defining $\kappa(x) = Kx$, the determination of $\bar{x}$ needs to be carried out by searching for a point $\bar{x} \in \mathbb{R}^3$, such that

$$\mathbf{0} \in \{A\bar{x} + B\mathcal{K}[\mathrm{q} \circ \kappa](\bar{x})\}.$$

On the other hand, for every $\bar{x} \in \mathbb{R}^3$, thanks to [97, Theorem 1], since rank $K = 2$, one has

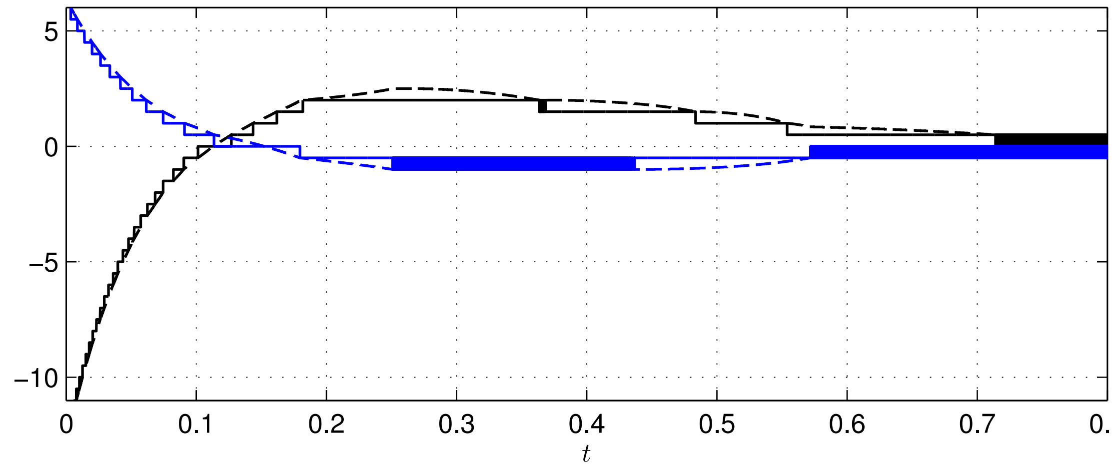
$$\mathcal{K}[\mathrm{q} \circ \kappa](\bar{x}) = \mathcal{K}[\mathrm{q}](K\bar{x}).$$

Moreover, due to the decentralized structure of the function $u \mapsto \mathrm{q}(u)$, from Proposition 1.3 it follows

$$\mathcal{K}[\mathrm{q} \circ \kappa](\bar{x}) = \bigtimes_{i=1}^{2} \mathcal{K}[\mathrm{q}](K_{(i)}\bar{x}).$$

Therefore, a necessary and sufficient condition for a point $\bar{x} \in \mathbb{R}^3$ to be a Krasovskii equilibrium for the closed-loop system is that

$$\mathbf{0} \in \{A\bar{x} + B \bigtimes_{i=1}^{2} \mathcal{K}[\mathrm{q}](K_{(i)}\bar{x})\}.$$



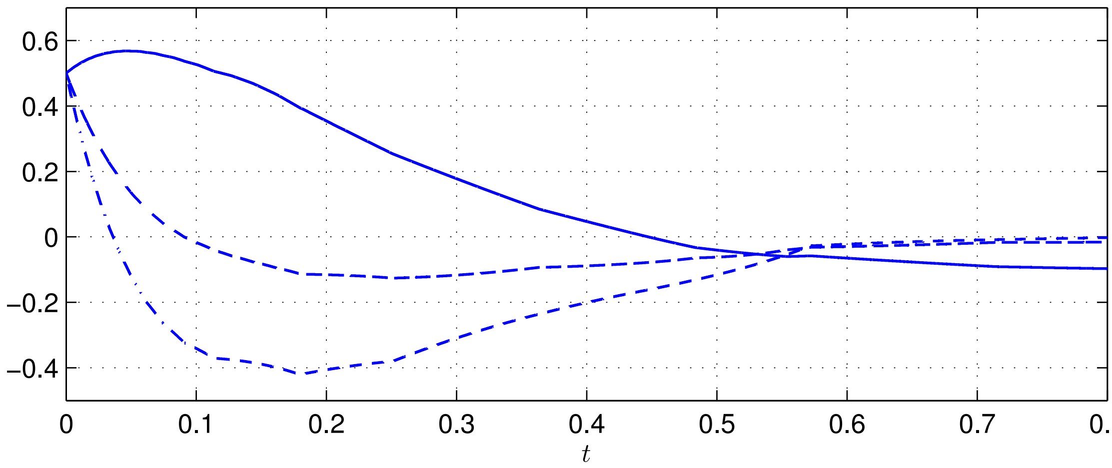


Figure 2.6: The evolution of the closed-loop system from $x_0 = (0.5, 0.5, 0.5)$: Above the control inputs ($\mathrm{q}(u_1)$ (solid-black), $\mathrm{q}(u_2)$ (solid-blue), and the two quantization-free inputs ($K_{(1)}x(t)$ (dashed-black), $K_{(2)}x(t)$ (dashed-blue). Below the closed-loop states: x_1 (solid), x_2 (dashed), x_3 (dashed-dotted). The solutions are obtained by integrating the closed-loop model via an Euler method with time step 10^{-4}.

Now, if one restricts the search to the points $\bar{x}$ such that $K\bar{x} = [\Delta \ -\Delta]^{\mathsf{T}}$, in view of the definition of the function $\mathrm{q}(\cdot)$ given in (1.17), the latter relation turns in

$$\mathbf{0} \in \left\{ \bar{x} \in \mathbb{R}^2 \colon A\bar{x} + B \begin{bmatrix} \delta_1 \\ \delta_2 \end{bmatrix} \Delta, (\delta_1, \delta_2) \in \{[0,1] \times [-1,0]\} \right\}.$$

Therefore, $\bar{x}$ needs to satisfy

$$\begin{cases} A\bar{x} = -B \begin{bmatrix} \delta_1 \\ \delta_2 \end{bmatrix} \Delta \\ K\bar{x} = \begin{bmatrix} \Delta \\ -\Delta \end{bmatrix} \end{cases}$$

and this is possible if and only if

$$\operatorname{rank}\left(\begin{bmatrix} A \\ K \end{bmatrix}\right) = \operatorname{rank}\left(\begin{bmatrix} A & -B\begin{bmatrix}\delta_1 \\ \delta_2\end{bmatrix} \\ K & \begin{bmatrix} 1 \\ -1 \end{bmatrix} \end{bmatrix}\right).$$

for some $(\delta_1, \delta_2) \in \{[0,1] \times [-1,0]\}$. In particular, it turns out that the latter condition is verified for $\delta_1 \approx 0.48856, \delta_2 \approx 0.21977$, which, in turn yields $\bar{x} = (-0.12, 0.018, -0.014)$. Notice that, $\bar{x}$ is the only Krasovskii equilibrium belonging to the surface $\mathcal{S}_1 := \{x \in \mathbb{R}^3 : Kx = [\Delta \ -\Delta]^\mathsf{T}\}$ for the closed-loop system. Analogous considerations allow to compute the other equilibrium point satisfying $K\bar{x}_2 = [-\Delta \ \Delta]^\mathsf{T}$, specifically $\bar{x}_2 = -\bar{x}$. *A posteriori* of these calculations, let us focus on Figure 2.7, that is essentially a closed-up of Figure 2.5, and in which the two computed equilibrium points are represented. The figures shows, both the accuracy of the above arguments in foreseeing the behavior of the closed-loop system, in terms of Krasovskii solutions, and the accuracy provided by Euler integration, that succeeds in capturing the peculiar behaviors due to the discontinuity introduced by the quantizer.

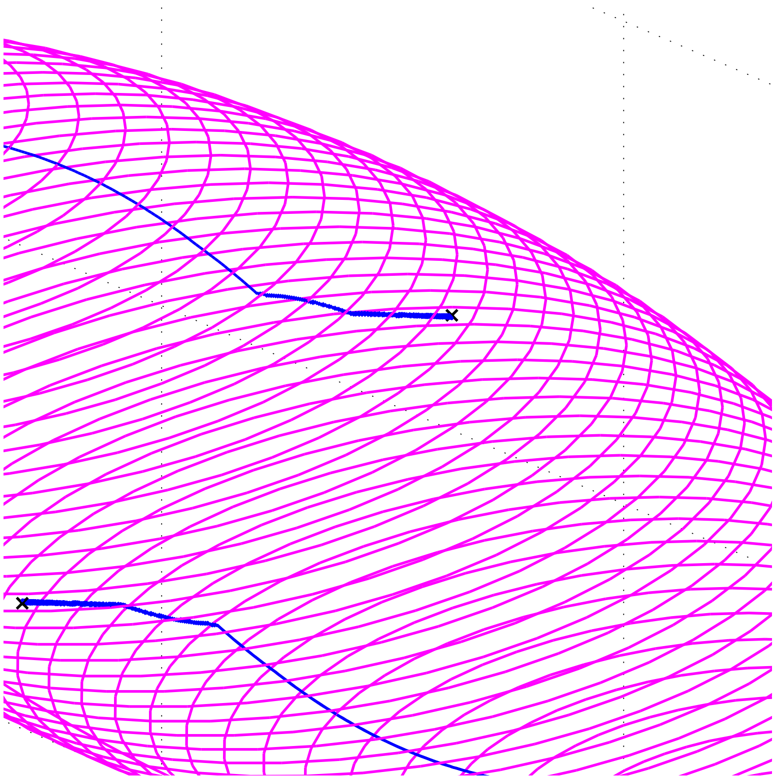


Figure 2.7: Closed-loop trajectories approaching the two equilibrium points ('x'). The solutions are obtained by integrating the closed-loop model via an Euler method with time step 10^{-4}

2.3 Sensor Quantization

2.3.1 Preliminary Results

Consider the following continuous-time linear system with quantized measured state

$$\begin{cases} \dot{x} = Ax + Bu \\ u = K\,\mathrm{q}(x) \end{cases} \tag{2.53}$$

where $x \in \mathbb{R}^n$, $u \in \mathbb{R}^m$, are respectively the state, the input of the system. A, B, K are real matrices of suitable dimensions, and $\mathrm{q}(\cdot)$ is the uniform quantizer defined in (1.17) having as a quantization error bound $\Delta > 0$. As in the previous section, by introducing the function Ψ defined in (2.2), the closed-loop system can be rewritten as

$$\dot{x} = (A + BK)x + BK\Psi(x). \tag{2.54}$$

Therefore, with the aim of considering the Krasovskii solutions to (2.54), let us define

$$\begin{aligned} Z\colon \mathbb{R}^n &\to \mathbb{R}^n \\ x &\mapsto (A + BK)x + BK\Psi(x) \end{aligned} \tag{2.55a}$$

we consider the solutions to the following differential inclusion.

$$\dot{x} \in \mathcal{K}[Z](x). \tag{2.55b}$$

By retracing the steps performed in the actuator quantization case, we provide a first result characterizing the behavior of the closed-loop system (2.54) in terms of its Krasovskii solutions.

Theorem 2.2. Let A, B, K be matrices of adequate dimensions, such that $A + BK$ is Hurwitz. Then there exists a compact set $\mathcal{A} \subset \mathbb{R}^n$, containing the origin, which is UGAS for (2.55).

Proof. The proof follows the same steps shown in the proof of Theorem 2.1, and then it is omitted. ■

Also in this case, we want to provide constructive tractable conditions for the search of the set $\mathcal{A} \subset \mathbb{R}^n$, whose existence is ensured by Theorem 2.2. Therefore, in the sequel, the same apparatus presented for the actuator quantization case is considered for the case of interest of this section.

2.3.2 Stability Analysis

Problem 2.5. (Stability analysis) Let A, B, K be matrices of adequate dimensions, such that $A + BK$ is Hurwitz. Determine a compact set $\mathcal{A} \subset \mathbb{R}^n$ containing the origin, such that

$\mathcal{A}$ is UGAS for system (2.55).

The next result, which essentially parallels Proposition 2.1, gives a first sufficient condition to solve Problem 2.5.

Proposition 2.8. *If there exist $P \in \mathcal{S}_+^n$, $S_1, S_2 \in \mathcal{D}_+^n$, and a positive scalar τ such that*

$$\begin{bmatrix} \mathrm{He}(P(A+BK)) + \tau P & PBK - S_2 \\ \bullet & -S_1 - 2S_2 \end{bmatrix} < \mathbf{0} \tag{2.56}$$

$$\mathrm{trace}(S_1)\Delta^2 - \tau \leq 0 \tag{2.57}$$

then,

$$\mathcal{A} = \mathcal{E}(P) \tag{2.58}$$

solves Problem 2.5.

Proof. The proof retraces the same steps performed in the proof of Proposition 2.1. For every $x \in \mathbb{R}^n$, consider the following quadratic function $V(x) = x^\mathsf{T} P x$. Following the ideas presented in the proof of Theorem 2.1, we want to prove that there exists a positive real scalar β such that

$$\langle \nabla V(x), w \rangle \leq -\beta V(x) \qquad \forall x \in \mathbb{R}^n \setminus \mathrm{Int}\mathcal{A}, w \in \mathcal{K}[Z](x). \tag{2.59}$$

As the above relation is analogous to (2.8) in the proof of Theorem 2.1, establishing (2.59) suffices to show that the set $\mathcal{A}$ in (2.58) is UGAS for (2.55). By S-procedure arguments, (2.59) can be verified by showing that for every $x \in \mathbb{R}^n$, there exists a positive real scalar τ such that

$$\langle \nabla V(x), w \rangle - \tau(1 - x^\mathsf{T} P x) \leq -\beta V(x) \qquad \forall w \in \mathcal{K}[Z](x). \tag{2.60}$$

On the other hand, via Proposition 1.1, for every $w \in \mathcal{K}[Z](x)$, there exists $v \in \mathcal{K}[\Psi](x)$, such that $w = (A + BK)x + BKv$. Then, still by S-procedure arguments and according to Lemma 2.2, (2.60) is ensured by proving that for each $x \in \mathbb{R}^n$, and for each $v \in \mathbb{R}^n$,

$$\begin{aligned} &\langle \nabla V(x), (A+BK)x + BKv \rangle - \tau(1 - x^\mathsf{T} P x) - v^\mathsf{T} S_1 v \\ &+ \mathrm{trace}(S_1)\Delta^2 - 2v^\mathsf{T} S_2 (v + x) \leq -\beta V(x). \end{aligned}$$

By straightforward calculations, the left-hand side of the above relation can be rewritten as follows

$$\begin{bmatrix} x \\ v \end{bmatrix}^T \begin{bmatrix} \mathrm{He}(P(A+BK)) + \tau P & PBK - S_2 \\ \bullet & -S_1 - 2S_2 \end{bmatrix} \begin{bmatrix} x \\ v \end{bmatrix} + \mathrm{trace}(S_1)\Delta^2 - \tau. \tag{2.61}$$

Thus in view of (2.56) and (2.57), it follows that there exists a small enough positive scalar γ such that for every $x \in \mathbb{R}^n \setminus \mathrm{Int}\mathcal{A}, w \in \mathcal{K}[Z](x)$, one has $\langle \nabla V(x), w \rangle \leq -\gamma x^\mathsf{T} x$. Then, since for every $x \in \mathbb{R}^n$, $V(x) \leq \lambda_{\max}(P) x^\mathsf{T} x$, by setting $\beta = \frac{\gamma}{\lambda_{\max}(P)}$ gives (2.60), and this finishes the proof. ∎

Also in this case, the feasibility of the conditions given in Proposition 2.8 is ensured under

Assumption 1.2. In particular, let us consider the following result, whose proof is essentially based on the ingredients exploited in the proof of Proposition 2.2.

Proposition 2.9. *Let $K \in \mathbb{R}^{m\times n}$ such that $A + BK$ is Hurwitz. Then, there exists $(\tau, P, S_1, S_2) \in \mathbb{R}_{>0} \times \mathcal{S}_+^n \times \mathcal{D}_+^n \times \mathcal{D}_+^n$ satisfying (2.56) and (2.57).* □

2.3.3 Controller Design

As already done in the actuator quantization case, also in this case, we want to tackle the controller design problem for the closed-loop system (2.55). Essentially, assuming that the gain K has to be designed, we want to derive tractable constructive conditions characterizing the solutions to the problem formalized as follows

Problem 2.6. (Controller design) Let A, B be matrices of adequate dimensions. Determine a gain $K \in \mathbb{R}^{m\times n}$, and a compact set $\mathcal{A} \subset \mathbb{R}^n$ containing the origin, such that $\mathcal{A}$ is UGAS for system (2.55).

Clearly Proposition 2.8 provides a first condition to solve Problem 2.6. However, due to products between unknown variables, a direct exploitation of the conditions given by Proposition (2.8) to solve the controller design problem appears unlikely, and then further work is needed. Nevertheless, differently from (2.14), applying similar strategies as the ones shown in Proposition 2.3 does not allow to alleviate the bilinear term PBK (and its transpose) appearing in (2.14). In particular, if one attempts to alleviate this term by means of standard techniques (congruence transformations, and invertible changes of variables), the resulting condition reveals to be still nonlinear and presenting more involved nonlinearities as trilinear terms. On the other hand, via the use of the projection lemma, (see, *e.g.*, [99]), one can derive a condition equivalent to (2.56), which is linear in the variable P defining the set $\mathcal{A}$, and bilinear with respect to the controller gain and some additional variables. This condition is proposed in the result given next.

Proposition 2.10. *Let $P \in \mathcal{S}_+^n, S_1, S_2 \in \mathcal{D}_+^n, K \in \mathbb{R}^{m\times n}$, and $\tau \in \mathbb{R}_{>0}$. The satisfaction of*

$$\begin{bmatrix} \mathrm{He}(P(A+BK)) + \tau P & PBK - S_2 \\ \bullet & -S_1 - 2S_2 \end{bmatrix} < \mathbf{0} \tag{2.62}$$

is equivalent to the feasibility of

$$\begin{bmatrix} -\mathrm{He}(X_1) & P - X_2 + X_1^{\mathsf{T}}(A+BK) & X_1^{\mathsf{T}} BK \\ \bullet & \mathrm{He}(X_2^{\mathsf{T}}(A+BK)) + \tau P & X_2^{\mathsf{T}} BK - S_2 \\ \bullet & \bullet & -S_1 - 2S_2 \end{bmatrix} < \mathbf{0} \tag{2.63}$$

with respect to $X_1, X_2 \in \mathbb{R}^{n\times n}$.

Proof. The proof follows the same lines of the one of Corollary 2.1. In particular, from

Proposition 2.8, notice that (2.56) can be rewritten as $\mathcal{W}^\mathsf{T}\mathcal{Q}\mathcal{W} < \mathbf{0}$, where:

$$\mathcal{W} = \begin{bmatrix} A+BK & BK \\ \mathrm{I} & \mathbf{0} \\ \mathbf{0} & \mathrm{I} \end{bmatrix}, \mathcal{Q} = \begin{bmatrix} \mathbf{0} & P & \mathbf{0} \\ \bullet & \tau P & -S_2 \\ \bullet & \bullet & -S_1 - 2S_2 \end{bmatrix}.$$

Moreover, being S_1 and S_2 positive definite, $\mathcal{U}^\mathsf{T}\mathcal{Q}\mathcal{U} < \mathbf{0}$, with $\mathcal{U}^\mathsf{T} = \begin{bmatrix} \mathbf{0} & \mathbf{0} & \mathrm{I} \end{bmatrix}$, is obviously satisfied. Thus, by the projection lemma; see [99], the satisfaction of (2.56), whenever S_1 and S_2 are required to be positive definite, is equivalent to find a matrix $\mathcal{X}$ such that

$$\mathcal{Q} + \mathcal{W}^{\mathsf{T}\perp}_r \mathcal{X}\mathcal{U}^\perp_r + \mathcal{U}^{\mathsf{T}\perp}_r \mathcal{X}^\mathsf{T}\mathcal{W}^\perp_r < \mathbf{0} \tag{2.64}$$

where, $\mathcal{U}^\perp_r$ and $\mathcal{W}^\perp_r$ are some matrices having as rows a basis of the row-null space, respectively, of $\mathcal{U}$ and $\mathcal{W}$. Now, by selecting $\mathcal{U}^\perp_r = \begin{bmatrix} \mathrm{I}_{2n} & \mathbf{0}_{2n\times p} \end{bmatrix}$ and $\mathcal{W}^\perp_r = \begin{bmatrix} -\mathrm{I} & A+BK & BK \end{bmatrix}$ and by partitioning $\mathcal{X} = \begin{bmatrix} X_1 & X_2 \end{bmatrix}$, where $X_1, X_2 \in \mathbb{R}^{n\times n}$, from (2.64) one gets

$$\begin{bmatrix} -\operatorname{He}(X_1) & P - X_2 + X_1^\mathsf{T}(A+BK) & X_1^\mathsf{T}BK \\ \bullet & \operatorname{He}(X_2^\mathsf{T}(A+BK)) + \tau P & X_2^\mathsf{T}BK - S_2 \\ \bullet & \bullet & -S_1 - 2S_2 \end{bmatrix} < \mathbf{0} \tag{2.65}$$

which is (2.63) and this finishes the proof. ■

As already mentioned, the advantage offered by the above result is twofold. On the one hand, there is no trilinear term. On the other hand, the matrix P defining the set $\mathcal{A}$ appears linearly. This fact enables to build an iterative relaxation procedure that allows to solve the controller design problem, this aspect is presented in the next subsection.

2.3.4 Optimization Issues

Concerning the optimization aspects in the solution to Problem 2.5 and Problem 2.6, analogous considerations as the ones presented for the actuator quantization case hold in this case. In particular, the optimization problems to address in this setting can be formulated as follows:

Problem 2.7 (Stability)**.** Let A, B, K be matrices of adequate dimensions. Determine $P \in \mathcal{S}^n_+$, such that $\mathcal{E}(P)$ is UGAS for system (2.55), and it is minimized with respect to some criterion.

Problem 2.8 (Stabilization)**.** Let A, B be matrices of adequate dimensions. Determine a gain $K \in \mathbb{R}^{m\times n}$, and $P \in \mathcal{S}^n_+$, such that $\mathcal{E}(P)$ is UGAS for system (2.55), and it is minimized with respect to some criterion.

Obviously, since the sets whose size needs to be minimized are still ellipsoidal sets, the size criteria that can be considered in Problem 2.7, and Problem 2.8 are the same illustrated for the actuator quantization case.

As far as concerns Problem 2.7, as long as τ is fixed, (2.56) is linear in the decision variables. Then, a direct generalization of Algorithm 2.1 allows to solve Problem 2.7 in a convex setting. Such an algorithm is given next

Algorithm 2.3 Stability analysis

Input: Matrices A, B, K, scalars $\Delta > 0$, a convex objective function $\mathcal{M}_a$, and a tolerance $\rho > 0$.

Initialization: Let $\mathcal{R}(A + BK) := \{|\Re(\lambda)|\colon \lambda \in \text{spec}(A + BK)\}$, select $\tau = 2 \times 0.99 \min \mathcal{R}(A + BK)$,

Iteration

Step 1:

Solve the following convex optimization problem over LMIs

$$\begin{aligned} \underset{S_1,S_2,P}{\text{minimize}} \quad & \mathcal{M}_a(P) \\ \text{s.t.} \quad & S_1, S_2 \in \mathcal{D}_+^n, P \in \mathcal{S}_+^n \\ & \begin{bmatrix} \text{He}(P(A+BK)) + \tau P & PBK - S_2 \\ \bullet & -S_1 - 2S_2 \end{bmatrix} < \mathbf{0} \\ & \text{trace}(S_1)\Delta^2 - \tau \leq 0 \end{aligned}$$

Pick the sub-optimal solution $(\overline{P}, \overline{S}_1, \overline{S}_2)$. Store the obtained solution: $\mathcal{M}_{a\star}^{(k)} \leftarrow \mathcal{M}_a(\overline{P})$, $P_\star^{(k)} \leftarrow \overline{P}$.

$k \leftarrow k + 1$

Step 2:

Decrease τ of ρ, *i.e.*, $\tau \leftarrow \tau - \rho$

Until $\tau > 0$.

Step 3: $k_{\max} \leftarrow k$, select $k^\star = \underset{k \in \{1,2,k_{\max}\}}{\text{argmin}} \{\mathcal{M}_{a\star}^{(k)}\}$

Output: $P = P_\star^{(k^\star)}$.

Clearly the same considerations pointed in Remark 2.4 holds also for the above algorithm.

As mentioned before, the solution to Problem 2.8, due to nonlinearities affecting condition (2.56) is much more involved, and requires a suitable strategy. In particular, inspired by [5], we propose the following iterative algorithm to derive a suboptimal solution to Problem 2.8.

Algorithm 2.4 Controller design

Input: Matrices A, B, scalar $\Delta > 0$, a convex objective function $\mathcal{M}_s$, and a desired tolerance $\rho > 0$.
Initialization: Select $\overline{K}$ such that $\overline{A}_{cl} = A + B\overline{K}$ is Hurwitz. Let $\mathcal{R}(\overline{A}_{cl}) \coloneqq \{|\Re(\lambda)| \colon \lambda \in \operatorname{spec}(\overline{A}_{cl})\}$, then select $\overline{\tau} = 2 \times 0.99 \min, \mathcal{R}(\overline{A}_{cl})$ and build a grid of positive values $\mathcal{G}_\tau$ such that $\max \mathcal{G}_\tau = \overline{\tau}$ (ensures the feasibility of the resulting optimization problems).
Iteration *Step 1:* Given $\overline{K}$ from the previous step, solve the following convex optimization problem over LMIs by selecting τ over $\mathcal{G}_\tau$.

$$\begin{aligned} \underset{S_1,S_2,P,X_1,X_2}{\text{minimize}} \quad & \mathcal{M}_s(P) \\ \text{s.t.} \quad & S_1, S_2 \in \mathcal{D}_+^n, P \in \mathcal{S}_+^n \\ & \begin{bmatrix} -\operatorname{He}(X_1) & P - X_2 + X_1^\mathsf{T}(A + B\overline{K}) & X_1^\mathsf{T} B\overline{K} \\ \bullet & \operatorname{He}(X_2^\mathsf{T}(A + B\overline{K})) + \tau P & X_2^\mathsf{T} B\overline{K} - S_2 \\ \bullet & \bullet & -S_1 - 2S_2 \end{bmatrix} < \mathbf{0} \\ & \operatorname{trace}(S_1)\Delta^2 - \tau \leq 0 \end{aligned} \tag{2.66}$$

Pick the suboptimal solution obtained and set $\overline{X}_1 = X_1, \overline{X}_2 = X_2$ for the next step.
Step 2: Given $\overline{X}_1, \overline{X}_2$ from the previous step, solve the following convex optimization problem over LMIs by selecting τ over $\mathcal{G}_\tau$.

$$\begin{aligned} \underset{S_1,S_2,P,K}{\text{minimize}} \quad & \mathcal{M}_s(P) \\ \text{s.t.} \quad & S_1, S_2 \in \mathcal{D}_+^n, P \in \mathcal{S}_+^n \\ & \begin{bmatrix} -\operatorname{He}(\overline{X}_1) & P - \overline{X}_2 + \overline{X}_1^\mathsf{T}(A + BK) & \overline{X}_1^\mathsf{T} BK \\ \bullet & \operatorname{He}(\overline{X}_2^\mathsf{T}(A + BK)) + \tau P & \overline{X}_2^\mathsf{T} BK - S_2 \\ \bullet & \bullet & -S_1 - 2S_2 \end{bmatrix} < \mathbf{0} \\ & \operatorname{trace}(S_1)\Delta^2 - \tau \leq 0. \end{aligned} \tag{2.67}$$

Set $\overline{K} = K$, for the next step.
Determine the closed-loop matrix $A + B\overline{K}$ and set $\overline{\tau} = 2 \times 0.99 \min \mathcal{R}(A + B\overline{K})$. Build a grid of positive values $\mathcal{G}_\tau$ such that $\bar{\tau} = \max \mathcal{G}_\tau$, and $\tau^\star \in \mathcal{G}_\tau$, (notice that necessarily $\tau^\star \leq \bar{\tau}$. Including $\tau^\star$ in $\mathcal{G}_\tau$ ensures the feasibility at the next step).
Until $\mathcal{M}_s$ does not decrease below ρ over three consecutive steps.
Output: $\overline{K}, P$.

Remark 2.10. Proposition 2.10 plays a determinant role in the development of the above exposed algorithm. In fact, the introduction of the slack variables X_1, X_2 enables to treat P as a decision variable at each step of the algorithm, without adding any additional conservatism; recall that the feasibility of (2.64) is equivalent to the one of (2.56). Notice that, by exploiting directly Proposition 2.8, due to the bilinear terms involving the matrix P and the controller gain K, if one would retrace the strategy proposed in Algorithm 2.4, then one needs to alternatively fix either P or K, preventing from treating P as a decision variable at each step. This obviously has a dramatic impact on the achievable suboptimal solutions to Problem 2.8.

The above algorithm presents some interesting properties that render its utilization in practice quite convenient. In particular, notice that at each iteration, both (2.66) and (2.67) are always feasible. Indeed, during the first iteration, since A_{cl} is Hurwitz, the feasibility of (2.66) is ensured by Proposition 2.9. To see that also at each other iteration the considered optimization problem are always feasible, consider the following arguments.

For the $j-th$ iteration, denote the value of the matrix P, respectively, at the exit of step 1 and of step 2 as $\overline{P}_j^{(1)}$ and $\overline{P}_j^{(2)}$.

[From step 1 to step 2] Obviously step 2 is always feasible, indeed keeping the same gain $\overline{K}$ from the previous step yields a feasible solution and moreover $\mathcal{M}_s(\overline{P}_j^{(2)}) \leq \mathcal{M}_s(\overline{P}_j^{(1)})$.

[From step 2 to step 1] The feasibility of (2.66) is ensured by following same arguments illustrated for the other case. Moreover, $\mathcal{M}_s(\overline{P}_{j+1}^{(1)}) \leq \mathcal{M}_s(\overline{P}_j^{(2)})$. Notice also that by assuming $\mathcal{M}_s(P) \geq 0$ over the feasible set of (2.67) (this assumption is certainly verified for the trace criterion previously illustrated and can be fulfilled for the logdet criterion by considering for the stopping criterion $-\det(P)$ which is positive on $\mathcal{S}_+^n$ and monotonically related to $-\operatorname{logdet}(P)$), the above mentioned monotonicity property guarantees that the sequence $\{\mathcal{M}_s(\overline{P}_j^{(2)}\}_{j=1}^{\infty}$ converges. Therefore, for any positive ρ, the algorithm terminates in a finite number of iterations.

As for the actuator quantization case, one may add (in step 2) further constraints to limit the norm of the controller gain. In particular, since in step 2 the controller gain K is a decision variable, for any positive κ, considering the following constraint (linear in K)

$$\begin{bmatrix} \kappa^2 \mathrm{I}_n & K \\ \bullet & \mathrm{I}_m \end{bmatrix} \leq \mathbf{0} \tag{2.68}$$

ensures that $\|K\| \leq \kappa$.

Remark 2.11. Although Algorithm 2.4 provides a numerically tractable solution to Problem 2.8, one should be aware that the initialization stage plays a relevant role in the final result. In particular, from different initializations the algorithm may converge to different solutions. On the other hand, determining the most efficient initialization seems a nontrivial problem.

As far as concerns the convex criterion $\mathcal{M}_s$ to adopt in Algorithm 2.4, both the determinant based criterion and the trace criterion as it was presented in (2.33) represent valuable choices that leads to computationally tractable procedures.

2.3.5 Numerical Examples

Example 2.4. Consider the static state feedback control system with quantized sensor from [49], that is defined by the following data:

$$A = \begin{bmatrix} 0 & 1 \\ 0.5 & 0.5 \end{bmatrix}, B = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \Delta = 1, K = \begin{bmatrix} -0.3491 & -0.7022 \end{bmatrix}.$$

By selecting as convex criterion $\mathcal{M}_s(P) = -\log\det(P)$, by solving Problem 2.7 via Algorithm 2.3, with a tolerance $\rho = 0.001$, yields

$$P = \begin{bmatrix} 2605570.22255 & -2605570.2217 \\ -2605570.2217 & 2605570.22332 \end{bmatrix}$$

Figure 2.8 reports the set $\mathcal{E}(P)$, along with some closed-loop solutions. Notice that

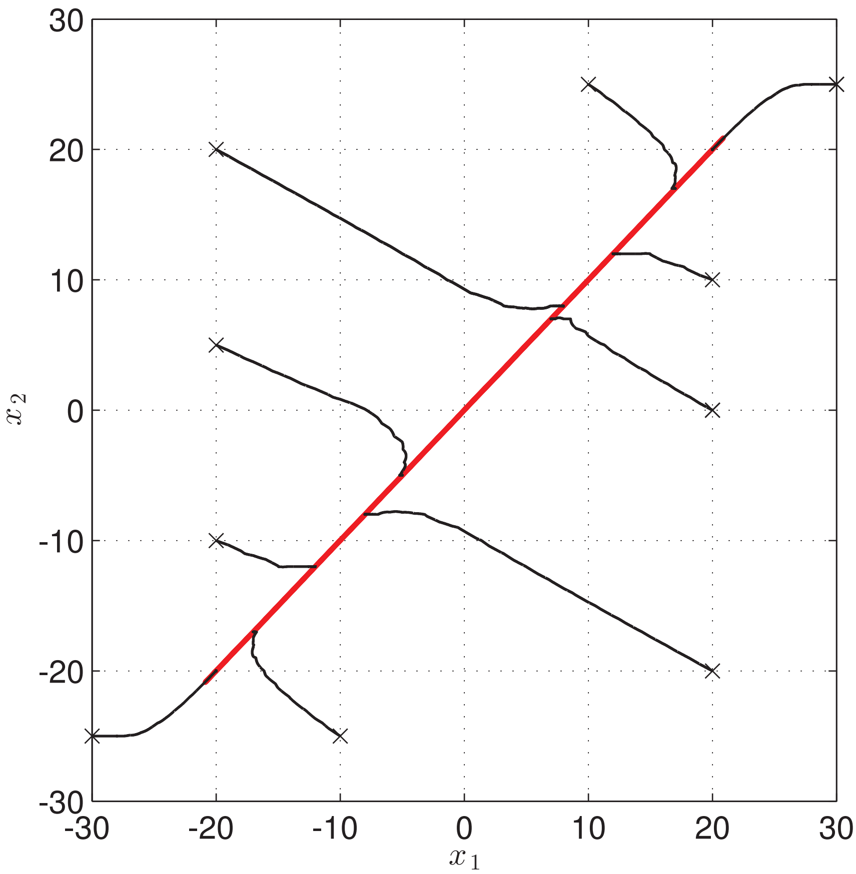


Figure 2.8: The set $\mathcal{E}(P)$ (red), some closed-loop trajectories (black). Solutions are obtained by integrating the closed-loop model with an Euler method with time step 10^{-4}.

$$\text{spec}(P) = \{0.001148, 5.211 \times 10^6\}$$

ensuring that $P > \mathbf{0}$. This huge difference between such eigenvalues is due to the shape of the set $\mathcal{E}(P)$ represented in Figure 2.8, which is nearly a segment.

Figure 2.8 points out that the solution to the Problem 2.7 via the proposed algorithm provides a very satisfactory characterization of the actual behavior of the closed-loop system. In particular, simulations suggest that the closed-loop trajectories converges toward the set $\mathcal{S} := \{(x_1, x_2) \in \mathbb{R}^2 \colon x_1 = x_2\} \cap \mathbb{Z}^2$. Clearly, such a set is not connected, therefore being $\mathcal{A}$ necessarily connected (in fact convex), it can only provide an overapproximation of $\mathcal{S}$, which seems quite tight in this case.

Fostered by the above arguments, one may wonder whether the points belonging to $\mathcal{S}$ are equilibrium points for the closed-loop system. However, such points, except for the origin, cannot be equilibria in a classical sense. Indeed, let us assume that there exists a classical equilibrium $\bar{x} \in \mathcal{S}$. Then, it has to be

$$\begin{aligned} &\bar{x}_1 = \bar{x}_2 = \bar{k} \\ &A\bar{x} + BK\mathbf{1}_2\bar{k} = 0 \end{aligned}$$

for some $k \in \mathbb{Z}$. That is

$$(A + BK)\mathbf{1}_2\bar{k} = 0$$

but the latter, being $A + BK$ Hurwitz, is obviously satisfied only for $\bar{k} = 0$. That said, the search of the equilibrium points into the set $\mathcal{S}$ needs to be performed by looking at Krasovskii equilibria. Similarly to Example 2.3, we seek for each point $\bar{x} \in \mathbb{R}^2$, such that

$$\mathbf{0} \in \{A\bar{x} + BK\mathcal{K}[\mathrm{q}](\bar{x})\}.$$

On the other hand, for every $x \in \mathbb{R}^2$, thanks to Proposition 1.3, one has

$$\mathcal{K}[\mathrm{q}](\bar{x}) = \bigtimes_{i=1}^{2} \mathcal{K}[\mathrm{q}](\bar{x}_{(i)})$$

Therefore, it follows that a point $\bar{x} \in \mathbb{R}^2$ is a Krasovskii equilibrium for the closed-loop system if and only if

$$\mathbf{0} \in \{A\bar{x} + BK \bigtimes_{i=1}^{2} \mathcal{K}[\mathrm{q}](\bar{x}_{(i)})\}.$$

Now, if one restricts the search to the points $\bar{x}$ such that $\bar{x} = k\mathbf{1}_2$, for some $k \in \mathbb{N}$, in view of the definition of the function $\mathrm{q}(\cdot)$ given in (1.17), the latter relation turns into

$$\mathbf{0} \in \left\{ \bar{x} \in \mathbb{R}^2 \colon A\bar{x} + BK \begin{bmatrix} \delta_1 \\ \delta_2 \end{bmatrix} : (\delta_1, \delta_2) \in \{[k-1, k] \times [k-1, k]\} \right\}.$$

Therefore, by setting $\bar{x} = k\mathbf{1}_n$, $\bar{x}$ is a Krasovskii equilibrium for the closed-loop system if and only if, there exists $(\delta_1, \delta_2) \in \{[k-1, k] \times [k-1, k]$, such that

$$A\mathbf{1}_n k + BK \begin{bmatrix} \delta_1 \\ \delta_2 \end{bmatrix} = 0.$$

Since

$$BK = \begin{bmatrix} K_{(1)} & K_{(2)} \\ K_{(1)} & K_{(2)} \end{bmatrix}, A\mathbf{1}_n = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

the latter equality imposes that

$$K_{(1)}\delta_1 + K_{(2)}\delta_2 = -k.$$

Concluding, $\bar{x} = k\mathbf{1}_2$, with $k \in \mathbb{N}$ can be a Krasovskii equilibria for the closed-loop system if and only if the following polyhedral in $\mathbb{R}^2$ is nonempty

$$\begin{cases} K_{(1)}\delta_1 + K_{(2)}\delta_2 = -k \\ \delta_1 \leq k \\ \delta_1 \geq k-1 \\ \delta_2 \leq k \\ \delta_2 \geq k-1. \end{cases} \tag{2.69}$$

Moreover, by definition of the uniform quantizer (1.17) and Proposition 1.1, if

$$\mathbf{0} \in \{A\bar{x} + BK \bigtimes_{i=1}^{2} \mathcal{K}[\mathrm{q}](\bar{x}_{(i)})\},$$

then

$$\mathbf{0} \in \{-A\bar{x} + BK \bigtimes_{i=1}^{2} \mathcal{K}[\mathrm{q}](-\bar{x}_{(i)})\}$$

that is the equilibria are symmetric with respect to the origin.

Therefore, in practice, to determine if the points $\pm k\mathbf{1}_2$ are Krasovskii equilibria for the closed-loop system for some $k \in \mathbb{N}$, one can test, via standard linear programming algorithms, whether (2.69) is non-empty. In particular, by pursuing this approach, it turns out that for the given gain K, (2.69) is non-empty for k up to 20. This means that the only Krasovskii equilibria belonging to $\mathcal{S}$ for the closed-loop system are the points $\bar{x} = k\mathbf{1}_2$ with $k = \pm 1, \pm 2 \ldots, \pm 20$, which exactly matches the results presented in Figure 2.8. Specifically, Figure 2.9 emphasizes that the closed-loop system solutions approach the Krasovskii equilibria.

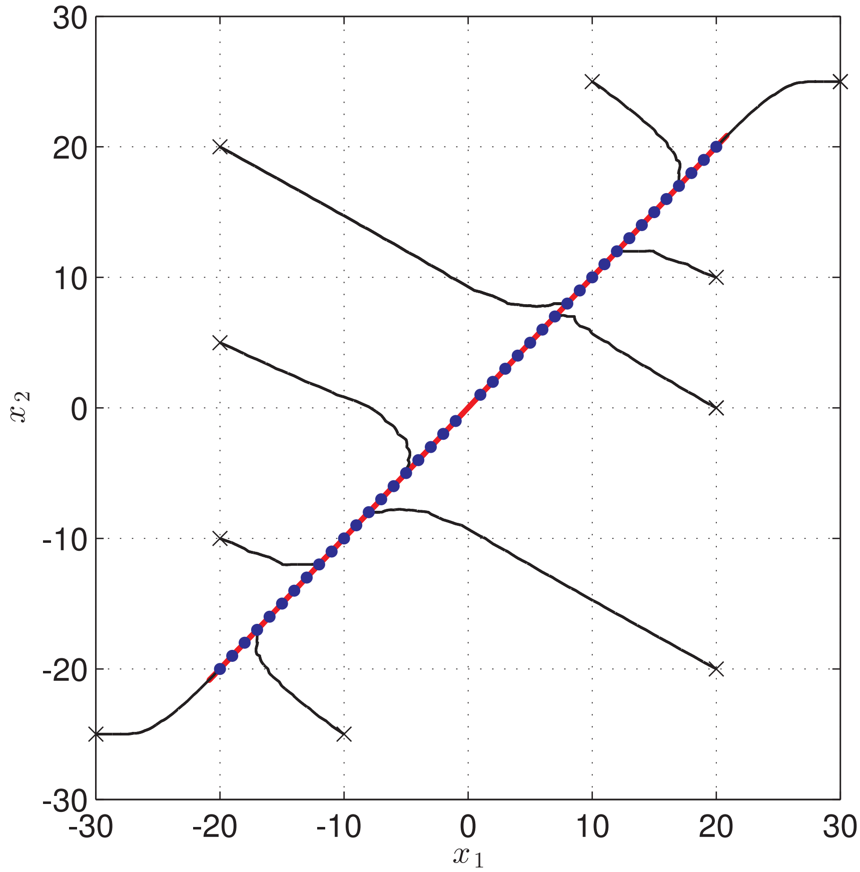


Figure 2.9: The set $\mathcal{E}(P)$ (red), some closed-loop trajectories (black), and the Krasovskii equilibria (blue bullets). Solutions are obtained by integrating the closed-loop model with an Euler method with time step 10^{-4}.

As pointed out above, the shape of the set $\mathcal{A}$ intrinsically leads to a matrix P which tends to be ill-conditioned. On the other hand, whenever having a good conditioning is a relevant matter, one may add an additional constraint in the considered optimization problem, so as to ensure a given condition number for the matrix P. Such constraint can easily integrated by means of additional LMI constraints; see [127], at the price of obtaining more conservative results. Indeed, limiting the condition number reflects on the shape of the resulting set $\mathcal{A}$. To show this fact in this example, for the considered closed-loop system, we solve Problem 2.7 via Algorithm 2.3, while considering an additional constraint aimed at ensuring a condition number for P less or equal than γ. Figure 2.10 reports the sets $\mathcal{A}$ obtained as above, whenever γ varies in a grid built upon the interval $[10, 5000]$. The figure shows that, as expected, the larger the condition number, the tighter the resulting set $\mathcal{A}$.

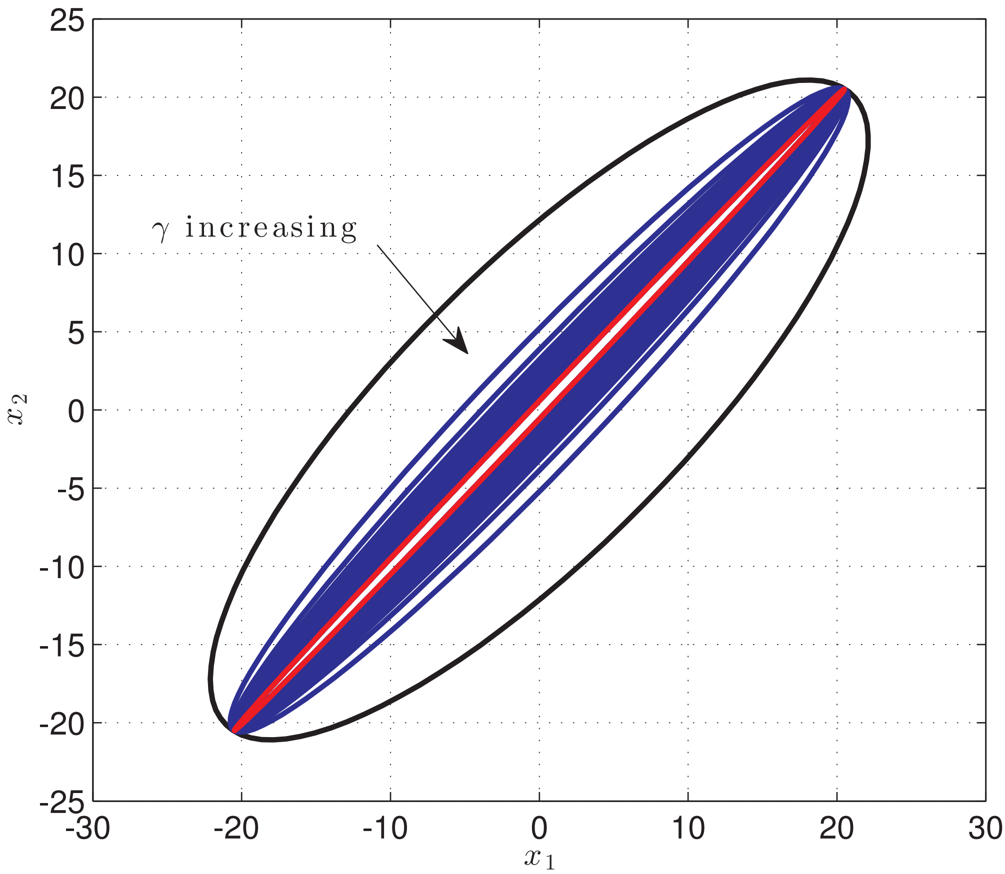


Figure 2.10: Different sets $\mathcal{A}$ obtained imposing a condition number for P less or equal than γ. $\gamma = 10$ (black), $\gamma = 5000$ (red), $\gamma \in (10, 5000)$ (blue), the sets shrink as γ increases.

Example 2.5. (A multi-input plant) Consider the again the example from [2] for which

$$A = \begin{bmatrix} -0.5 & 1.5 & 4 \\ 4.3 & 6 & 5 \\ 3.2 & 6.8 & 7.2 \end{bmatrix}, B = \begin{bmatrix} -0.7 & -1.3 \\ 0 & -4.3 \\ 0.8 & -1.5 \end{bmatrix},$$

and assume in this case, that the measured state is quantized via the uniform quantizer (1.17) with $\Delta = 0.5$. We want to solve Problem 2.8 via Algorithm 2.4, by using the trace criterion as presented in (2.33). In particular, let $N \in \mathcal{S}_+^n$, by requiring that

$$\begin{bmatrix} N & \mathrm{I} \\ \bullet & P \end{bmatrix} \geq \mathbf{0}$$

we want to minimize trace(N). To initialize the algorithm, we use three different stabilizing gains. The first one

$$K_0 = \begin{bmatrix} 0.0380 & 0.1751 & -0.8551 \\ 3.8514 & 3.8400 & 9.5510 \end{bmatrix}$$

is borrowed directly from [2]. The second one

$$K_1 = \begin{bmatrix} -0.71 & 1.9 & -27 \\ 4.3 & 4.1 & 4.3 \end{bmatrix}$$

comes from Example 2.3, and finally the third one,

$$K_2 = \begin{bmatrix} -0.11527 & -0.28207 & -1.2449 \\ 2.4835 & 4.2519 & 6.2107 \end{bmatrix}$$

is the gain issued from the solution of an LQR problem on the pair (A, B), with $Q = \mathrm{I}_3$, and $R = \mathrm{I}_2$. For all these three initializations, the tolerance for the algorithm is $\rho = 10^{-4}$. Figure 2.11 shows the evolution of trace(N) over the number of iterations for the three proposed initializations. Surprisingly, although the algorithm does not ensures convergence toward the optimal solution, and the initialization are quite different of each other, the algorithm provides three solutions giving nearly the same value of the objective. This shows that, at least for the matter of this specific example, the initialization stage is not excessively crucial, though it may impact the computational burden: the number of iterations might increase depending on the initialization, *e.g.*, for the third initialization the number of iterations is almost twice as much as the number of iterations occurring for the second initialization. In Table 2.1 the different outputs of the algorithm are reported for the three considered initializations. As shown in Table 2.1, the first and the third initialization provides quite similar results also in terms of controller gain and the matrix P defining the set $\mathcal{A} = \mathcal{E}(P)$ solving the controller design problem.

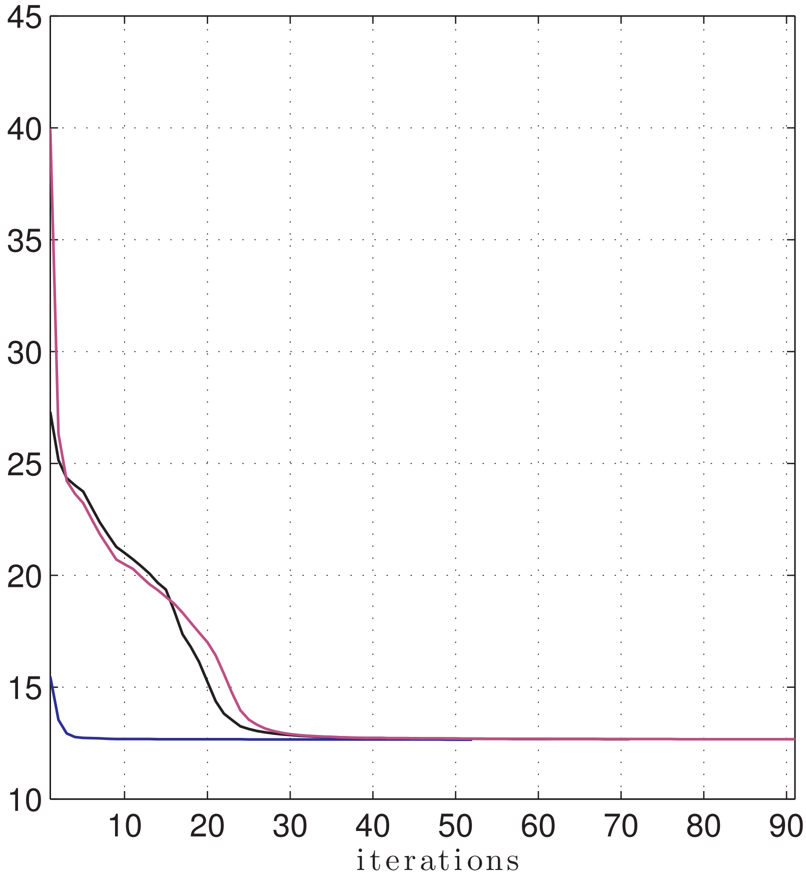


Figure 2.11: Objective function versus the number of iterations. First initialization (magenta), second initialization (blue), third initialization (black).

Initialization	trace(N)	K	P	Iterations
1	12.673	$\begin{bmatrix} -3.3894 & -0.37076 & -25.912 \\ 4.4094 & 0.57293 & 16.594 \end{bmatrix}$	$\begin{bmatrix} 0.99 & 0.3 & -0.53 \\ 0.3 & 0.37 & -0.6 \\ -0.53 & -0.6 & 1.4 \end{bmatrix}$	71
2	12.653	$\begin{bmatrix} -8.4812 & -0.1307 & -30.799 \\ 6.3613 & 1.4677 & 13.504 \end{bmatrix}$	$\begin{bmatrix} 0.59 & 0.18 & -0.13 \\ 0.18 & 0.42 & -0.55 \\ -0.13 & -0.55 & 1.1 \end{bmatrix}$	52
3	12.669	$\begin{bmatrix} -3.4714 & -0.39353 & -25.839 \\ 4.3721 & 0.60222 & 16.29 \end{bmatrix}$	$\begin{bmatrix} 0.97 & 0.3 & -0.53 \\ 0.3 & 0.37 & -0.61 \\ -0.53 & -0.61 & 1.4 \end{bmatrix}$	91

Table 2.1: The different outputs of Algorithm 2.4 for the three different initializations.

2.4 Comments and Conclusion

In this chapter, we addressed the state feedback control problem for linear systems in the presence of either actuator quantization or sensor quantization, in terms of Krasovskii solutions. In this setting, first we shown that the asymptotic stability of the quantization-free closed-loop system (in both the considered schemes) ensures the existence of an ellipsoidal set $\mathcal{A}$, which UGAS for the closed-loop system (in terms of its Krasovskii solutions). Then, with the aim of pursuing a constructive approach, thanks to some novel sector conditions for the uniform quantizer, we turn the search of the set $\mathcal{A}$ into the feasibility problem of certain matrix inequalities. Moreover, we shown that this approach is lossless, in the sense that under the asymptotic stability of the quantization-free closed-loop system, the derived matrix inequalities are always feasible.

As a second step, we addressed the stabilization problem for the same class of systems. In this context, the considered problem consists of deriving some conditions enabling the simultaneous search of a linear static state feedback controller, and a compact set $\mathcal{A}$ containing the origin, such that the resulting closed-loop system has the set $\mathcal{A}$ UGAS. Such a problem is solved by suitably transforming the matrix inequalities derived for the stability problem in more advantageous fashions.

Building on the derived conditions, some algorithms based on convex optimization over LMIs are proposed to effectively solve the considered problems, while providing (sub)optimal solutions with respect to convenient objectives. Finally, the effectiveness of the proposed methodology is shown in some examples. These examples, not only provide a benchmark to test the proposed apparatus from a numerical standpoint, but also point out the complexity hidden behind quantized control systems.

Although, the proposed methodology is tailored to the uniform quantizer defined in (1.17), the framework is quite flexible to envision extensions to other type of quantizers. For instance, the extension to the uniform quantizer considered in [22] is quite straightforward. In particular, give $\Delta > 0$, such a quantizer is defined for each $u \in \mathbb{R}^n$ as

$$\mathrm{q}(u) = \left\lfloor \frac{u}{\Delta} + \frac{1}{2} \right\rfloor \Delta$$

where the above operators are considered component-wise. Therefore, analogously to the case considered in this chapter, define for each $u \in \mathbb{R}^n$ the function $\Gamma(u) = \mathrm{q}(u) - u$. As pointed in Figure 2.12, it can be readily shown that for each $S_1, S_2 \in \mathcal{D}^n_+$ and for each $u \in \mathbb{R}^n$, the following conditions hold for the function Γ

$$\begin{aligned}
&\Gamma(u)^{\mathsf{T}} S_1 \Gamma(u)^{\mathsf{T}} - \mathrm{trace}(S_1)\frac{\Delta^2}{4} \leq 0 \\
&(\Gamma - u)^{\mathsf{T}} S_2 (\Gamma + u) \leq 0.
\end{aligned}$$

Hence, the methodology illustrated in this chapter can be extended to deal with the quantizer

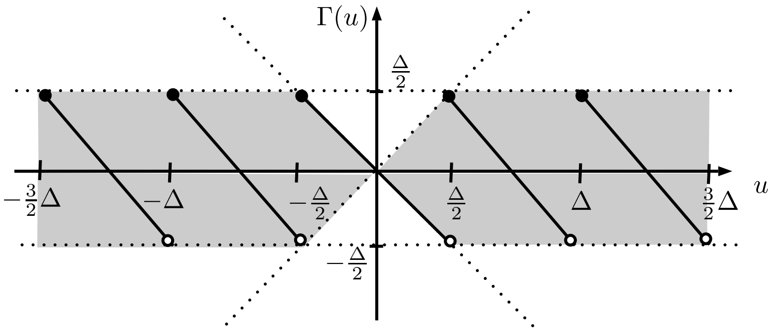


Figure 2.12: The function Γ, for the scalar case and its sector.

adopted in [22] with few extra work.

In the case of combined sensor and actuator quantization, the stability problem could be addressed in the same manner, though this case requires a special care in view of the nested discontinuous nonlinearity issued by the combined effect of the two quantizers. Concerning the design problem, due to the difficulties encountered even for the simpler case of the mere sensor quantization, a design procedure in this case appears intricate. Nevertheless, we will show how to solve this problem later in the next chapter, by employing a dynamical output feedback controller.

Also, the work presented in this chapter assumes that the quantizer has an unbounded range. However, if one considers a finite range for the quantizer, using ideas from [82], the proposed methodology can be still adapted, providing local results. In particular, as far as concerns the actuator quantization case, let us assume that the uniform quantizer defined in (1.17) has a finite range M, that is for each $u \in \mathbb{R}^m$, $q(u) = (q(u_1), q(u_2), \ldots, q(u_m))$, and

$$q_{(i)}(u) = \begin{cases} \Delta \operatorname{sign}(u_{(i)}) \left\lfloor \frac{|u_{(i)}|}{\Delta} \right\rfloor & u_{(i)} \in [-M, M] \\ M & \text{otherwise.} \end{cases}$$

Consider the set $\mathcal{A} = \mathcal{E}(P)$ obtained from the solution to Problem 1. Building from this set, pick $\eta \in (0, 1)$, and consider the set $\mathcal{A}_M = \mathcal{E}(P, \eta) := \{x \in \mathbb{R}^n : x^\mathsf{T} P x \leq \eta\}$, then $\mathcal{A} \subset \mathcal{A}_M$. Define the set $\mathcal{S}(K, M) := \{x \in \mathbb{R}^m : |Kx| \preceq M\}$. If there exists $\eta > 1$, such that $\mathcal{A}_M \subset \mathcal{S}(K, M)$, then all the arguments presented in the proof of Theorem 2.2 are still valid inside the set $\mathcal{A}_M$, hence local uniform asymptotic stability of the set $\mathcal{A}$ can be established directly. From this observation, it appears obvious that all the result presented in this chapter can be extended to derive conditions ensuring that the set $\mathcal{A}$ is locally uniformly asymptotically stable for the closed-loop system in the presence of finite range uniform quantization, without no much modifications. Analogous considerations hold for the sensor quantization case. Clearly, in this setting more involved optimization problems could be considered. For instance, the minimization of the set $\mathcal{A}$ could be coupled with the maximization of the set $\mathcal{A}_M$, still with respect to adequate size criteria. An interesting point to address in the actuator quantization case, in the presence of finite range quantizers, regards the design of the gain K to simultaneously enlarge $\mathcal{A}_M$ and shrink $\mathcal{A}$.

Although this approach enables to solve the two considered problems in the presence of finite range quantization, the results obtained by restraining the set $\mathcal{A}_M$ to be contained in the set $\mathcal{S}(K, M)$ may be conservative. Indeed, as for the case of saturating systems, one may enable the quantizer to saturate while ensuring the well behavior of the closed-loop system. Clearly, this approach requires a dedicate strategy. For instance, by using ideas from [117], one may model the finite range quantizer, as the composition of an infinite range quantizer and a standard saturation operator, say $\operatorname{sat}(\cdot)$. Pursuing this approach enables to blend the techniques proposed by the literature of saturated systems, with the techniques presented in this chapter. On the other hand, one should be aware that handling the closed-loop system in terms of Krasovskii solutions, in this case requires further work. Just to give an hint about the difficulties encountered in this case, consider that the closed-loop system, in this case, reads

$$\dot{x} = Ax + B\operatorname{sat}(\mathrm{q}(Kx)).$$

Therefore, the differential inclusion issuing from the Krasovskii regularization of the right-hand side of the above expression gives

$$\dot{x} \in Ax + B\mathcal{K}[\operatorname{sat} \circ \mathrm{q} \circ K](x)$$

where with an abuse of notation, we denoted K the linear operator issuing from the matrix K, *i.e.*, the function $x \mapsto Kx$. Obviously, the latter needs to be suitably worked out to distinguish the effect of the two nonlinearities[3]. This is work is currently part of our research activity.

Concerning the actuator quantization case, another interesting aspect consists of considering the effect induced by replacing the actual state with an estimate provided by an asymptotic observer, whenever the plant state is not fully accessible. In particular, let us consider the following plant

$$\begin{cases} \dot{x} = Ax + Bu \\ u = \mathrm{q}(Kx) \\ y = Cx \end{cases} \tag{2.70}$$

where $y \in \mathbb{R}^p$ is the measured output. In particular, as the plant dynamics are linear, we consider the following full-order Luenberger state observer; [88]

$$\dot{\hat{x}} = A\hat{x} + Bu + L(y - C\hat{x}) \tag{2.71}$$

where $\hat{x} \in \mathbb{R}^n$ is the estimate of the plant state x provided by the observer, and $L \in \mathbb{R}^{n\times p}$ is the observer gain to be designed. Building on the estimate provided by (2.71), we consider the following control law

$$u = K\hat{x} \tag{2.72}$$

[3]Notice that, since $\mathrm{q} \circ K$ is discontinuous, Proposition 1.2 does not provide any viable strategy to build an overapproximation for $\mathcal{K}[\operatorname{sat} \circ \mathrm{q} \circ K](x)$.

where $K \in \mathbb{R}^{m\times n}$ is the controller gain to be designed. By means of the latter control law, the dynamics of the closed-loop system (2.70)-(2.71) can be written as

$$\begin{cases} \dot{x} = Ax + B\,\mathrm{q}(K\hat{x}) \\ \dot{\hat{x}} = A\hat{x} + B\,\mathrm{q}(K\hat{x}) + LC(x-\hat{x}) \end{cases} \tag{2.73}$$

Since the state $\hat{x}$ can be seen as an estimate of x, by defining the estimation error $\varepsilon = x-\hat{x}$, the dynamics of (2.73) can be rewritten in a more convenient fashion in the (x,ε) coordinates. In particular, by defining the following invertible change of variables

$$\begin{bmatrix} x \\ \varepsilon \end{bmatrix} = \begin{bmatrix} \mathrm{I} & \mathbf{0} \\ \mathrm{I} & -\mathrm{I} \end{bmatrix} \begin{bmatrix} x \\ \hat{x} \end{bmatrix}$$

by taking as vector state $\tilde{x} = (x,\varepsilon) \in \mathbb{R}^{2n}$, and by defining, for each $u \in \mathbb{R}^m$, the function $\Psi(u) = \mathrm{q}(u) - u$, the closed-loop system (2.73), in the new coordinate turns into

$$\dot{\tilde{x}} = \underbrace{\begin{bmatrix} A+BK & -BK \\ \mathbf{0} & A-LC \end{bmatrix}}_{A_c} \tilde{x} + \underbrace{\begin{bmatrix} B \\ \mathbf{0} \end{bmatrix}}_{B_c} \Psi\Big(\underbrace{\begin{bmatrix} K & -K \end{bmatrix}}_{C_c} \tilde{x}\Big). \tag{2.74}$$

Therefore, with the aim of considering Krasovskii solutions to (2.74), define

$$\begin{aligned} X\colon \mathbb{R}^{2n} &\to \mathbb{R}^{2n} \\ \tilde{x} &\mapsto A_c\tilde{x} + B_c\Psi(C_c\tilde{x}) \end{aligned} \tag{2.75}$$

we consider the solutions to the following differential inclusion

$$\dot{\tilde{x}} \in \mathcal{K}[X](\tilde{x}). \tag{2.76}$$

At this stage, notice that (2.75)-(2.76) inherits some notable properties by the upper triangular structure of (2.74). In particular, it is not difficult to show that whenever $A-LC$ is Hurwitz, the set $\mathbb{R}^n \times \{0\}$ is UGAS for (2.75)-(2.76). Moreover, notice that every solution $\tilde{\phi} = (\phi_x, \phi_\varepsilon)$ to the restriction of (2.76) to the set $\mathbb{R}^n \times \{0\}$, *i.e.*,

$$\dot{\tilde{x}} \in \mathcal{X}(\tilde{x}) \tag{2.77}$$

where, for each $\tilde{x} \in \mathbb{R}^n$, $\mathcal{X} := \mathcal{K}[X](\tilde{x}) \cap (\mathbb{R}^n \times \{0\})$, is such that ϕ_x is a solution to (2.4). From this analysis, it is straightforward to show that the fulfillment of the conditions provided by Proposition 2.1, along with the Hurwitzness of the matrix $A-LC$ ensures that the set $\mathcal{E}(P) \times \{0\}$, where P is defined in Proposition 2.1, is UGAS for (2.75)-(2.76). A formal proof of this result arises from the application of [56, Corollary 7.24.], for the simpler case of differential inclusions, (an example pertaining to the cascade of two nonlinear systems is shown in [55]). Beyond the discussed properties arising from Hurwitzness of $A-LC$, and the upper triangular structure of (2.4), the key ingredients of the proof are that solutions to (2.4) are bounded for every initial condition (this is mainly due to the linearity of the plant dynamics), and that $\mathcal{K}[X]$ is convex-valued, outer semicontinuous, locally bounded, and $\operatorname{dom} \mathcal{K}[X] = \mathbb{R}^{2n}$. These arguments show that the apparatus built in this chapter for the design of a static state feedback stabilizer controller in the presence of actuator quantization can be considered also when the state is not fully accessible and replaced by an estimate generated through a Luenberger state observer.

Although the extension to the case of partial measurements in the presence of actuator quantization is trivial, the same extension in the case of sensor quantization is nontrivial and requires further work. This is the object of the subsequent chapter.

This page is intentionally left blank.

3

QUANTIZED DYNAMIC OUTPUT FEEDBACK STABILIZATION

"Scientific research is one of the most exciting and rewarding of occupations."

– Frederick Sanger

3.1 Introduction

THIS chapter pertains to output feedback stabilization of linear plants subject to sensor and actuator uniform quantization. In particular, we design a dynamic output feedback controller to achieve closed-loop UGAS of an ellipsoidal set. As a first stage, we consider that only the plant output is gathered by a uniformly quantized sensor. In this setting, we first provide a general result turning the stabilization problem into the feasibility problem to certain matrix inequalities. Then, we propose a methodology based on convex optimization over LMIs to design the stabilizing controller. As a second stage, we extend the approach mentioned above to tackle the same stabilization problem for linear plants subject to simultaneous sensor and actuator quantization. Finally, the proposed methodology is shown in some examples. Some of the results presented in this chapter can be found in [37, 38].

3.2 Sensor Quantization

3.2.1 Preliminary Results and Problem Statement

Consider the following continuous-time linear system with quantized sensor

$$\begin{cases} \dot{x} = Ax + Bu \\ y_m = \mathrm{q}(Cx) \end{cases} \tag{3.1}$$

where $x \in \mathbb{R}^n$, $u \in \mathbb{R}^m$, $y_m \in \mathbb{R}^p$, are respectively the state, the input, and the measured output of the plant. While $(A, B, C) \in \mathbb{R}^{n\times n} \times \mathbb{R}^{n\times m} \times \mathbb{R}^{p\times n}$, and $\mathrm{q}(\cdot)$ is the uniform quantizer defined in (1.17) having as a quantization error bound $\Delta > 0$. We want to design the following plant-order dynamic output feedback stabilizing controller for (3.1)

$$\begin{cases} \dot{x}_c = A_c x_c + B_c u_c \\ y_c = C_c x_c + D_c u_c \end{cases} \tag{3.2}$$

where $x_c \in \mathbb{R}^n$ is the controller state, $y_c \in \mathbb{R}^m$ is the controller output, $u_c \in \mathbb{R}^p$ is the controller input.

$$(A_c, B_c, C_c, D_c) \in \mathbb{R}^{n\times n} \times \mathbb{R}^{n\times p} \times \mathbb{R}^{m\times n} \times \mathbb{R}^{m\times p} \tag{3.3}$$

are real matrices to be designed. Interconnecting plant (3.1) with controller (3.2), *i.e.*, setting $u = y_c$, $u_c = y_m$, yields the following dynamics for the closed-loop system

$$\begin{cases} \dot{x} = Ax + BC_c x_c + BD_c\,\mathrm{q}(Cx) \\ \dot{x}_c = A_c x_c + B_c\,\mathrm{q}(Cx). \end{cases} \tag{3.4}$$

Therefore, as in Chapter 2, by defining the function

$$\begin{aligned} \Psi \colon \mathbb{R}^p &\to \mathbb{R}^p \\ z &\mapsto \mathrm{q}(z) - z \end{aligned} \tag{3.5}$$

by taking as vector state $\tilde{x} = (x, x_c) \in \mathbb{R}^{2n}$, and by defining the matrices

$$\widetilde{A} = \begin{bmatrix} A + BD_c C & BC_c \\ B_c C & A_c \end{bmatrix}, \widetilde{B} = \begin{bmatrix} BD_c \\ B_c \end{bmatrix}, \widetilde{C} = \begin{bmatrix} C & \mathbf{0} \end{bmatrix} \tag{3.6}$$

(3.4) can be rewritten as

$$\dot{\tilde{x}} = \widetilde{A}\tilde{x} + \widetilde{B}\Psi\left(\widetilde{C}\tilde{x}\right). \tag{3.7}$$

Since the function Ψ is discontinuous, the right-hand side of (3.7) is a discontinuous function of the state. Therefore, as done in the previous chapter, we want to focus on the Krasovskii solutions to such a system. Notice that, as for the other considered cases, in view of the local boundedness of the right-hand side of (3.7), for every $\tilde{x}_0 \in \mathbb{R}^{2n}$, there exists at least a

Krasovskii solution φ to (3.7) with $\varphi(0) = \tilde{x}_0$; (see Chapter 1). In particular, define

$$\begin{aligned} X \colon \mathbb{R}^{2n} &\to \mathbb{R}^{2n} \\ \tilde{x} &\mapsto \widetilde{A}\tilde{x} + \widetilde{B}\Psi(\widetilde{C}\tilde{x}) \end{aligned} \tag{3.8a}$$

we consider the solutions to the following differential inclusion

$$\dot{\tilde{x}} \in \mathcal{K}[X](\tilde{x}) \tag{3.8b}$$

where $\mathcal{K}[X](\tilde{x})$ represents the Krasovskii regularization of the function X; see Definition 1.2 on Page 14.

As pointed out on Page 25, the presence of the uniform quantizer, due to its deadzone effect, represents a real obstacle to the asymptotic stabilization of the origin of the closed-loop system. Specifically, if the matrix A is not Hurwitz, then the asymptotic stability of the origin for the closed-loop system (3.8) cannot be achieved via any choice of the controller (3.2). Nevertheless, also in this case, under suitable conditions on the quantization-free closed-loop system, system (3.8) manifests some interesting properties. In this sense, let us consider the following result that parallels Theorem 2.1.

Theorem 3.1. Let $A, B, C, A_c, B_c, C_c, D_c$ be matrices of adequate dimensions, such that $\widetilde{A}$ defined in (3.7) is Hurwitz. Then, there exists a compact set $\mathcal{A} \subset \mathbb{R}^{2n}$, containing the origin, which is UGAS for system (3.8).

Proof. The proof of the above result follows the same lines of the proof of Theorem 2.1. Thus, we provides the main steps of such proof below. In particular, under the considered hypothesis, we derive for (3.8) a relation like (2.8). Then, the proof directly follows from the arguments presented in the proof of Theorem 2.1.

For every $\tilde{x} \in \mathbb{R}^{2n}$, define $c(\tilde{x}) = \widetilde{C}\tilde{x}$. Since the function $\tilde{x} \mapsto \widetilde{A}\tilde{x}$ is continuous, by Proposition 1.1, for every $\tilde{x} \in \mathbb{R}^{2n}$, one has

$$\mathcal{K}[X](\tilde{x}) = \widetilde{A}\tilde{x} + \widetilde{B}\mathcal{K}[\Psi \circ c](\tilde{x}).$$

Since Ψ is locally bounded, (in fact bounded), thanks to [9, Lemma 1] it follows that, for every $\tilde{x} \in \mathbb{R}^{2n}$

$$\mathcal{K}[\Psi \circ c](\tilde{x}) = \operatorname{co}\left\{\lim \Psi(\widetilde{C}\tilde{x}_k) | \tilde{x}_k \to \tilde{x}\right\}$$

where $\tilde{x}_k$ is any sequence converging to $\tilde{x}$. Then, due to the bound shown in Chapter 2 on the function Ψ, it turns out that for each $\tilde{x} \in \mathbb{R}^{2n}$

$$\mathcal{K}[\Psi \circ c](\tilde{x}) \subseteq \mathbb{B}\sqrt{p}\Delta.$$

Therefore, for every $\tilde{x} \in \mathbb{R}^{2n}$, the following inclusion holds

$$\mathcal{K}[X](\tilde{x}) \subseteq \widetilde{A}\tilde{x} + \widetilde{B}\mathbb{B}\sqrt{p}\Delta. \tag{3.9}$$

Since $\widetilde{A}$ is Hurwitz, then there exist $P, Q \in \mathcal{S}_+^{2n}$, such that $\mathrm{He}(P\widetilde{A}) = -Q$. Building on this relation, for each $\tilde{x} \in \mathbb{R}^{2n}$, define the function $V(\tilde{x}) = \tilde{x}^\mathsf{T} P \tilde{x}$. Then, thanks to (3.9), for every $\tilde{x} \in \mathbb{R}^{2n}$, and any $f \in \mathcal{K}[X](\tilde{x})$

$$\langle \nabla V(x), f \rangle = -\tilde{x}^\mathsf{T} Q \tilde{x} + 2\tilde{x}^\mathsf{T} P \widetilde{B} \xi \leq -\lambda_{\min}(Q) \tilde{x}^\mathsf{T} \tilde{x} + 2\tilde{x}^\mathsf{T} P \widetilde{B} \xi$$

for some $\xi \in \mathbb{B}\sqrt{p}\Delta$. At this stage, by following the same arguments shown in the proof of Theorem 2.1, pick, $\theta \in (0, 1)$, then the compact set

$$\mathcal{A} = \left\{ \tilde{x} \in \mathbb{R}^{2n} \colon V(\tilde{x}) \leq \frac{4\lambda_{\max}(P)}{\lambda_{\min}^2(Q)\theta} \left\| \widetilde{B}^\mathsf{T} P^2 \widetilde{B} \right\| \Delta^2 p \right\}$$

is UGAS for (3.8), concluding the proof. ■

The above result shows that under the asymptotic stability of the quantization-free closed-loop system, there exists a compact set containing the origin, which is UGAS for the closed-loop system (3.8). On the one hand, Theorem 3.1 allows to select the more convenient notion of stability to consider in dealing with (3.8), and points out that if the controller is selected among the stabilizing controllers for the quantization free dynamics, then the set $\mathcal{A}$ is a sublevel set of a certain quadratic function. On the other hand, the above result gives rise to the same considerations addressed for Theorem 2.1. Indeed, Theorem 3.1 provides a coarse characterization of the behavior of (3.8), whose tightness dramatically depends on the choices of the controller and of the matrix P. Therefore, with the aim of designing the controller (3.2) to mitigate the effect induced by sensor quantization, the adoption of Theorem 3.1 is of any help. For this reason, as already done in the previous chapter, we pursue a constructive approach. Specifically, we derive computationally tractable conditions characterizing the solutions to the problem formalized as follows.

Problem 3.1. (Controller design) Let A, B, C be matrices of adequate dimensions. Determine matrices $(A_c, B_c, C_c, D_c) \in \mathbb{R}^{n \times n} \times \mathbb{R}^{n \times p} \times \mathbb{R}^{m \times n} \times \mathbb{R}^{m \times p}$ and a compact set $\mathcal{A} \subset \mathbb{R}^{2n}$ containing the origin, such that $\mathcal{A}$ is UGAS for system (3.8).

The solution to the above problem is the object of the remainder of this chapter. Specifically, in the sequel, by retracing the same approach carried out in the previous chapter, we present a complete apparatus to turn the solution to Problem 3.1 into the solution to certain matrix inequalities, while considering optimization aspects.

3.2.2 Sufficient Conditions

A first sufficient condition to solve Problem 3.1, and based on the sector conditions illustrated in Lemma 2.2, is given next.

Proposition 3.1. *If there exist $P \in \mathcal{S}_+^{2n}$, $S_1, S_2 \in \mathcal{D}_+^p$, A_c, B_c, C_c, D_c real matrices of ade-*

quate dimensions, and a positive scalar τ *such that*

$$\begin{bmatrix} \mathrm{He}(P\widetilde{A}) + \tau P & P\widetilde{B} - \widetilde{C}^\mathsf{T} S_2 \\ \bullet & -2S_2 - S_1 \end{bmatrix} < \mathbf{0} \tag{3.10}$$

$$\mathrm{trace}(S_1)\Delta^2 - \tau \leq 0 \tag{3.11}$$

where $\widetilde{A}, \widetilde{B}, \widetilde{C}$ *are defined in* (3.6)*. Then* A_c, B_c, C_c, D_c *and*

$$\mathcal{A} = \mathcal{E}(P) \tag{3.12}$$

solve Problem 3.1.

Proof. For every $\tilde{x} \in \mathbb{R}^{2n}$, consider the following quadratic function $V(\tilde{x}) = \tilde{x}^\mathsf{T} P \tilde{x}$. Following the ideas presented in the proof of Theorem 2.1, we want to prove that there exists a positive real scalar β such that

$$\langle \nabla V(\tilde{x}), w \rangle \leq -\beta V(\tilde{x}) \qquad \forall \tilde{x} \in \mathbb{R}^{2n} \setminus \mathrm{Int}\mathcal{A}, w \in \mathcal{K}[X](\tilde{x}). \tag{3.13}$$

As the above relation is analogous to (2.8) in the proof of Theorem 2.1, establishing (3.13) suffices to show that the set $\mathcal{A}$ in (3.12) is UGAS for system (3.8). By S-procedure arguments, (3.13) can be verified by showing that for every $\tilde{x} \in \mathbb{R}^{2n}$, there exists a positive real scalar τ such that

$$\langle \nabla V(\tilde{x}), w \rangle - \tau(1 - \tilde{x}^\mathsf{T} P \tilde{x}) \leq -\beta V(\tilde{x}) \qquad \forall w \in \mathcal{K}[X](\tilde{x}). \tag{3.14}$$

On the other hand, via Proposition 1.1 and Proposition 1.2, for every $w \in \mathcal{K}[X](\tilde{x})$, there exists $v \in \mathcal{K}[\Psi](\widetilde{C}\tilde{x})$, such that $w = \widetilde{A}\tilde{x} + \widetilde{B}v$. Then, still by S-procedure arguments and according to Lemma 2.2, (3.14) is ensured by proving that for each $\tilde{x} \in \mathbb{R}^{2n}$, and for each $v \in \mathbb{R}^p$,

$$\langle \nabla V(\tilde{x}), \widetilde{A}\tilde{x} + \widetilde{B}v \rangle - \tau(1 - \tilde{x}^\mathsf{T} P \tilde{x}) - v^\mathsf{T} S_1 v + \mathrm{trace}(S_1)\Delta^2 - 2v^\mathsf{T} S_2(v + \widetilde{C}\tilde{x}) \leq -\beta V(\tilde{x}). \tag{3.15}$$

By straightforward calculations, the left-hand side of the above relation can be rewritten as follows

$$\begin{bmatrix} \tilde{x} \\ v \end{bmatrix}^T \begin{bmatrix} \mathrm{He}(P\widetilde{A}) + \tau P & P\widetilde{B} - \widetilde{C}^\mathsf{T} S_2 \\ \bullet & -2S_2 - S_1 \end{bmatrix} \begin{bmatrix} \tilde{x} \\ v \end{bmatrix} + \mathrm{trace}(S_1)\Delta^2 - \tau. \tag{3.16}$$

Thus in view of (3.10) and (3.11), it follows that there exists a small enough positive scalar γ such that for every $\tilde{x} \in \mathbb{R}^{2n} \setminus \mathrm{Int}\mathcal{A}, w \in \mathcal{K}[X](\tilde{x})$, one has $\langle \nabla V(x), w \rangle \leq -\gamma \tilde{x}^\mathsf{T} \tilde{x}$. Then, since for every $\tilde{x} \in \mathbb{R}^{2n}$, $V(\tilde{x}) \leq \lambda_{\max}(P)\tilde{x}^\mathsf{T}\tilde{x}$, by setting $\beta = \frac{\gamma}{\lambda_{\max}(P)}$ gives (3.14), and this finishes the proof. ∎

The above result provides a sufficient condition to solve Problem 3.1. As in Chapter (2), Assumption 1.2 ensures the feasibility of conditions (3.10) and (3.11). This claim is formalized in the result given next.

Proposition 3.2. *Let* A, B, C *matrices such that Assumption 1.2 is satisfied. Then, there* *exist*

$$(\tau, P, S_1, S_2, A_c, B_c, C_c, D_c) \in \mathbb{R}_{>0} \times \mathcal{S}_+^{2n} \times \mathcal{D}_+^p \times \mathcal{D}_+^p \times \mathbb{R}^{n\times n} \times \mathbb{R}^{n\times p} \times \mathbb{R}^{m\times n} \times \mathbb{R}^{m\times p}$$

satisfying (3.10) *and* (3.11).

Proof. Notice that, from Assumption 1.2, there always exist A_c, B_c, C_c, D_c such that $\widetilde{A}$ is Hurwitz. Therefore, since (3.10) and (3.11) match, respectively, (2.14) and (2.15), by following the same arguments proposed in the proof of Proposition 2.2, the assert is proven. ■

The conditions provided by Proposition 3.1 turns the solution to Problem 3.1 into a feasibility problem to certain matrix inequalities. However, (3.10) is nonlinear in the decision variables, therefore, in general, solving Problem 3.1 by directly solving the feasibility problem associated to (3.10) and (3.11) appears unlikely from a numerical standpoint. To overcome this drawback, in the sequel we show two possible strategies to derive computationally tractable conditions from Proposition 3.1. The first strategy consists of performing a special choice for the controller parameters in (3.3) and for the matrix P in (3.10). Such choices spring from the selection of a linear observer-based controller. The second strategy instead consists of selecting a general output feedback dynamic controller and then capitalizing on existing results presented in the literature for the LMI-based design of output feedback dynamic controllers.

3.2.3 Controller design: Observer-based like Controller Design

The solution presented in this section builds on the following result.

Proposition 3.3. *If there exist* $P_1, P_2 \in \mathcal{S}_+^n$*,* $S_1, S_2 \in \mathcal{D}_+^p$*,* $K \in \mathbb{R}^{m\times n}$*,* $L \in \mathbb{R}^{n\times p}$*, and a positive scalar* τ *such that*

$$\begin{bmatrix} \mathrm{He}(P_1(A+BK)) + \tau P_1 & -P_1 BK & -C^{\mathsf{T}} S_2 \\ \bullet & \mathrm{He}(P_2(A-LC)) + \tau P_2 & -P_2 L \\ \bullet & \bullet & -2S_2 - S_1 \end{bmatrix} < \mathbf{0} \tag{3.17}$$

$$\mathrm{trace}(S_1)\Delta^2 - \tau \leq 0 \tag{3.18}$$

then,

$$A_c = A + BK - LC \tag{3.19}$$

$$B_c = L \tag{3.20}$$

$$C_c = K \tag{3.21}$$

$$D_c = \mathbf{0} \tag{3.22}$$

$$\mathcal{A} = \mathcal{E}\left(\begin{bmatrix} P_1 + P_2 & -P_2 \\ \bullet & P_2 \end{bmatrix}\right) \tag{3.23}$$

are a solution to Problem 3.1.

Proof. The proof of the above result is performed by showing that via congruence transformation and invertible change of variables, (3.10) turns into (3.17) for the choice of the controller given in (3.23) and a particular choice of the matrix P in (3.10) that is shown later. To this end, let us replace the controller parameters in (3.10) through the corresponding expressions given in (3.23). Via this step, $\widetilde{A}, \widetilde{B}$ in (3.6) turn into

$$\widetilde{A} = \begin{bmatrix} A & BK \\ LC & A + BK - LC \end{bmatrix}, \widetilde{B} = \begin{bmatrix} \mathbf{0} \\ L \end{bmatrix}. \tag{3.24}$$

Now, define

$$\Theta = \begin{bmatrix} \mathrm{I} & \mathbf{0} \\ \mathrm{I} & -\mathrm{I} \end{bmatrix}$$

and notice that since Θ is nonsingular (in fact $\Theta^{-1} = \Theta$), the satisfaction of (3.10) is equivalent to

$$\begin{bmatrix} \mathrm{He}(\Theta^{-\mathsf{T}} P \widetilde{A} \Theta^{-1}) + \tau \Theta^{-\mathsf{T}} P \Theta^{-1} & \Theta^{-\mathsf{T}} P \widetilde{B} - \Theta^{-\mathsf{T}} \widetilde{C}^{\mathsf{T}} S_2 \\ \bullet & -2S_2 - S_1 \end{bmatrix} < \mathbf{0} \tag{3.25}$$

which can be rewritten equivalently as follows:

$$\begin{bmatrix} \mathrm{He}(\Theta^{-\mathsf{T}} P \Theta^{-1} \Theta \widetilde{A} \Theta^{-1}) + \tau \Theta^{-\mathsf{T}} P \Theta^{-1} & \Theta^{-\mathsf{T}} P \widetilde{B} - \Theta^{-\mathsf{T}} \widetilde{C}^{\mathsf{T}} S_2 \\ \bullet & -2S_2 - S_1 \end{bmatrix} < \mathbf{0} \tag{3.26}$$

In particular, due to expression of $\Theta, \widetilde{A}, \widetilde{B}, \widetilde{C}$, by denoting

$$P = \begin{bmatrix} X & U \\ \bullet & \hat{X} \end{bmatrix},$$

one has:

$$\begin{aligned} \Theta^{-\mathsf{T}} P \Theta^{-1} &= \begin{bmatrix} X + \hat{X} + \mathrm{He}(U) & -\hat{X} - U \\ \bullet & \hat{X} \end{bmatrix} \\ \Theta \widetilde{A} \Theta^{-1} &= \begin{bmatrix} A + BK & -BK \\ \mathbf{0} & A - LC \end{bmatrix} \\ \Theta^{-\mathsf{T}} P \widetilde{B} &= \begin{bmatrix} (U + \hat{X})L \\ -\hat{X}L \end{bmatrix}. \end{aligned} \tag{3.27}$$

At this stage, select $U = -\hat{X}$, which gives

$$P = \begin{bmatrix} X & -\hat{X} \\ \bullet & \hat{X} \end{bmatrix} \tag{3.28}$$

according to this selection, from (3.27) one gets

$$\Theta^{-\mathsf{T}} P \Theta^{-1} = \begin{bmatrix} X - \hat{X} & \mathbf{0} \\ \bullet & \hat{X} \end{bmatrix}, \Theta^{-\mathsf{T}} P \widetilde{B} = \begin{bmatrix} \mathbf{0} \\ -\hat{X} L \end{bmatrix}. \tag{3.29}$$

To conclude, set $P_1 = X - \hat{X}$, $P_2 = \hat{X}$. Then, exploiting the latter change of variables and by plugging the expressions given in (3.29) into (3.27) yields (3.17). Therefore, the satisfaction of (3.17) and (3.18) implies the one of (3.10) and (3.11), whenever A_c, B_c, C_c, D_c are chosen as in (3.23) and

$$P = \begin{bmatrix} P_1 + P_2 & -P_2 \\ \bullet & P_2 \end{bmatrix}$$

which is symmetric and positive definite[1] due to $P_1 > \mathbf{0}$ and $P_2 > \mathbf{0}$, and this concludes the proof. ■

It is not difficult to realize that the choice of the controller, as long as the structure of the matrix Θ adopted to derive the above result actually build on a linear observer-based controller paradigm. Specifically, the considered controller is an observer-based controller, while Θ is the matrix associated to the classical change of variables leading to the closed-loop system represented in the (x, ε) coordinates, where ε represents the estimation error introduced by the observer. The selection of this controller stems from a few considerations. On the one hand, since the plant dynamics are linear, inspired by "certainty equivalence" principle illustrated in [83], it turns out that selecting an observer-based control revolving on a Luenberger observer seems the most natural choice to tackle the considered problem. On the other hand, Proposition 3.3 manifests two important features. The first one is that the provided result is lossless with respect to Theorem 3.1, in the sense that if there exist two gains K, L such that $A + BK$ and $A - LC$ are Hurwtiz, then the conditions provided by Proposition 3.3 are always feasible. Namely, Proposition 3.3 states a separation principle for the considered observer-based control architecture and for the stabilization objective pointed in Problem 3.1. The second one is that the considered result, by structuring the controller parameters, decreases the number of parameters to be designed and allows, through an adequate change of variables, to determine the gain L via convex optimization over LMIs, with the only caveat to make a choice for the gain K. The last shortcoming is quite common in the literature; see, *e.g.*, [79, 120]. These two properties are stated and formally proven as follows.

Fact 3.1. Let A, B, C, K, L be matrices such that $A + BK$ and $A - LC$ are Hurwitz. Then, there exists $(\tau, P_1, P_2, S_1, S_2) \in \mathbb{R}_{>0} \times \mathcal{S}_+^n \times \mathcal{S}_+^n \times \mathcal{D}_+^p \times \mathcal{D}_+^p$ satisfying (3.17) and (3.18).

Proof. The proof follows the lines of the proof of Proposition 2.2. Assume that there exists

[1] This claim can be readily proven by observing that the Schur complement of P is P_1.

$(\overline{\tau}, \overline{P_1}, \overline{P_2}, \overline{S_1}) \in \mathbb{R}_{>0} \times \mathcal{S}_+^n \times \mathcal{S}_+^n \times \mathcal{D}_+^p$ such that

$$\begin{bmatrix} \mathrm{He}(P_1(A+BK)) + \tau P_1 & -P_1 BK & \mathbf{0} \\ \bullet & \mathrm{He}(P_2(A-LC)) + \tau P_2 & -P_2 L \\ \bullet & \bullet & -S_1 \end{bmatrix} < \mathbf{0} \tag{3.30}$$

$$\mathrm{trace}(S_1)\Delta^2 - \tau \leq 0. \tag{3.31}$$

For each diagonal $S_2 \in \mathbb{R}^{p\times p}$, let us define

$$\mathcal{M}(S_2) := \begin{bmatrix} \mathrm{He}(\bar{P}_1(A+BK)) + \overline{\tau}\overline{P}_1 & -\overline{P}_1 BK & -C^{\mathsf{T}} S_2 \\ \bullet & \mathrm{He}(\bar{P}_2(A-\overline{L}C)) + \overline{\tau}\overline{P}_2 & -\bar{P}_2\bar{L} \\ \bullet & \bullet & -2S_2 - \bar{S}_1 \end{bmatrix}.$$

Since from (3.30) $\mathcal{M}(\mathbf{0}) < \mathbf{0}$ and $\mathcal{M}(S_2)$ depends continuously on the entries of S_2, there exists a small enough positive scalar δ, such that every $S_2 \in \mathcal{D}_+^p$ with $S_2 \leq \delta\mathrm{I}$ yields $\mathcal{M}(S_2) < \mathbf{0}$.

To conclude the proof, it suffices to show that whenever $A+BK$ and $A-LC$ are Hurwitz, there exists $(\overline{\tau}, \overline{P_1}, \overline{P_2}, \overline{S_1}) \in \mathbb{R}_{>0} \times \mathcal{S}_+^n \times \mathcal{S}_+^n \times \mathcal{D}_+^p$ satisfying (3.30) and (3.31). Define $A_k = A + BK$, $A_o = A - LC$, and let $\mathcal{R}(A_k) := \{|\Re(\lambda)| \colon \lambda \in \mathrm{spec}(A_k)\}$, $\mathcal{R}(A_o) := \{|\Re(\lambda)| \colon \lambda \in \mathrm{spec}(A_o)\}$. Notice that since A_k and A_o are Hurwitz, then $\mathcal{R}(A_k), \mathcal{R}(A_o) \subset \mathbb{R}_{>0}$. Pick $\overline{\tau} \in (0, 2\min\{\min\mathcal{R}(A_k), \min\mathcal{R}(A_o)\})$. Define, $\widetilde{A}_k = A_k + \frac{\overline{\tau}}{2}\mathrm{I}$, and $\widetilde{A}_o = A_o + \frac{\overline{\tau}}{2}\mathrm{I}$. Observe that, according to the selection considered for $\overline{\tau}$, $\widetilde{A}_k$ and $\widetilde{A}_o$ are Hurwitz. Select $\bar{S}_1 \in \mathcal{D}_+^p$, such that $\mathrm{trace}(\bar{S}_1)\Delta^2 - \overline{\tau} \leq 0$. By following these choices, the right-hand side of (3.30) reads

$$\begin{bmatrix} \mathrm{He}(P_1\widetilde{A}_k) & -P_1 BK & \mathbf{0} \\ \bullet & \mathrm{He}(P_2\widetilde{A}_o) & -P_2 L \\ \bullet & \bullet & -\bar{S}_1 \end{bmatrix}. \tag{3.32}$$

For any $\bar{Q}_2 \in \mathcal{S}_+^n$, pick the solution $\bar{W}_2 \in \mathcal{S}_+^n$ to the following matrix equality

$$\mathrm{He}(\tilde{A}_o\bar{W}_2) = -\bar{Q}_2 - L\bar{S}_1^{-1}L^{\mathsf{T}}$$

notice that such a solution always exists since $\tilde{A}_o$ is Hurwitz and $\bar{S}_1 \in \mathcal{D}_+^p$. For any $\bar{Q}_1 \in \mathcal{S}_+^n$, pick the solution $\bar{W}_1 \in \mathcal{S}_+^n$ to the following matrix equality

$$\mathrm{He}(\tilde{A}_k\bar{W}_1) = -\bar{Q}_1 - BK\bar{W}_2\bar{Q}_2^{-1}\bar{W}_2K^{\mathsf{T}}B^{\mathsf{T}}$$

still such a solution always exists since $\tilde{A}_k$ is Hurwitz and $\bar{Q}_2^{-1} \in \mathcal{S}_+^n$. Now, set in (3.30), $P_1 = \bar{W}_1^{-1}$, and $P_2 = \bar{W}_2^{-1}$. By following these choices, the right-hand side of (3.30) turns

$$\underbrace{\begin{bmatrix} \mathrm{He}(\bar{W}_1^{-1}\widetilde{A}_k) & -\bar{W}_1^{-1}BK & \mathbf{0} \\ \bullet & \mathrm{He}(\bar{W}_2^{-1}\widetilde{A}_o) & -\bar{W}_2^{-1}L \\ \bullet & \bullet & -\bar{S}_1 \end{bmatrix}}_{\mathscr{M}} \tag{3.33}$$

We want to show that $\mathscr{M}$ is negative definite. By pre-and-post multiplying $\mathscr{M}$ by $\mathrm{diag}(\bar{W}_1, \bar{W}_2, \mathrm{I})$, this is equivalent to show that

$$\begin{bmatrix} \mathrm{He}(\widetilde{A}_k \bar{W}_1) & -BK\bar{W}_2 & \mathbf{0} \\ \bullet & \mathrm{He}(\widetilde{A}_o \bar{W}_2) & -L \\ \bullet & \bullet & -\bar{S}_1 \end{bmatrix} < \mathbf{0}. \tag{3.34}$$

By Schur complement, since $\bar{S}_1 \in \mathcal{D}_+^p$, the latter relation holds if and only if

$$\begin{bmatrix} \mathrm{He}(\widetilde{A}_k \bar{W}_1) & -BK\bar{W}_2 \\ \bullet & \mathrm{He}(\widetilde{A}_o \bar{W}_2) + L\bar{S}_1^{-1}L^\mathsf{T} \end{bmatrix} < \mathbf{0} \tag{3.35}$$

which, due to the selection operated for $\bar{W}_1, \bar{W}_2$ turns in

$$\begin{bmatrix} -\bar{Q}_1 - BK\bar{W}_2\bar{Q}_2^{-1}\bar{W}_2K^\mathsf{T}B^\mathsf{T} & -BK\bar{W}_2 \\ \bullet & -\bar{Q}_2 \end{bmatrix} < \mathbf{0}. \tag{3.36}$$

By Schur complement, since $\bar{Q}_2 \in \mathcal{S}_+^n$, the latter is true if and only if

$$-\bar{Q}_1 - BK\bar{W}_2\bar{Q}_2^{-1}\bar{W}_2K^\mathsf{T}B^\mathsf{T} + BK\bar{W}_2\bar{Q}_2^{-1}\bar{W}_2K^\mathsf{T}B^\mathsf{T} = -\bar{Q}_1 < \mathbf{0}$$

which is obviously satisfied since $\bar{Q}_1 \in \mathcal{S}_+^n$. Therefore, $(\bar{\tau}, \bar{W}_1^{-1}, \bar{W}_2^{-1}, \bar{S}_1)$ satisfies (3.30) and (3.31), establishing the result. ■

Now, we illustrate the above mentioned change of variables allowing to partially linearize (3.17).

Corollary 3.1. If there exist $P_1, P_2 \in \mathcal{S}_+^n$, $S_1, S_2 \in \mathcal{D}_+^p$, $K \in \mathbb{R}^{m\times n}$, $J \in \mathbb{R}^{n\times p}$, and a positive scalar τ such that

$$\begin{bmatrix} \mathrm{He}(P_1(A+BK)) + \tau P_1 & -P_1BK & -C^\mathsf{T}S_2 \\ \bullet & \mathrm{He}(P_2A - JC) + \tau P_2 & -J \\ \bullet & \bullet & -2S_2 - S_1 \end{bmatrix} < \mathbf{0} \tag{3.37}$$

$$\mathrm{trace}(S_1)\Delta^2 - \tau \leq 0 \tag{3.38}$$

then,

$$A_c = A + BK - P_2^{-1}JC,\ B_c = P_2^{-1}J$$
$$C_c = K$$
$$D_c = \mathbf{0}$$

and

$$\mathcal{A} = \mathcal{E}\left(\begin{bmatrix} P_1 + P_2 & -P_2 \\ \bullet & P_2 \end{bmatrix}\right)$$

solve Problem 3.1.

Proof. The proof of the above result is straightforward. In particular, define the invertible change of variable $J = P_2 L$. Since the latter turns (3.17) into (3.37), the result is proven. ■

Optimization Issues

As already mentioned, in solving Problem 3.1, roughly speaking the main objective consists of designing the controller (3.2) to ensure that the solutions to (3.8) converge/stay sufficiently close to the origin. To this end, building on the conditions provided by Proposition 3.3, one can consider the following optimization problem.

Problem 3.2 (Observer-based stabilization)**.** Let A, B, C be matrices of adequate dimensions. Determine $K \in \mathbb{R}^{m\times n}$, $L \in \mathbb{R}^{n\times p}$, and $P_1, P_2 \in \mathcal{S}_+^n$, such that the set

$$\mathcal{E}\Bigg(\underbrace{\begin{bmatrix} P_1 + P_2 & -P_2 \\ \bullet & P_2 \end{bmatrix}}_{P}\Bigg) \tag{3.39}$$

is UGAS for system (3.8), and it is minimized with respect to some criterion.

As already illustrated in Chapter 2, the solution to the above optimization problem can be carried out by embedding the conditions provided by Proposition 3.3 into a suitable optimization scheme. To this end, an adequate measure of the set $\mathcal{E}(P)$ defined in (3.39) needs to be selected. As in Chapter 2, a first choice is to consider the volume of $\mathcal{E}(P)$ as a size criterion. In particular, with the aim of obtaining a convex optimization problem over LMIs, one can consider as a size criterion $-\log\det(P)$. In particular, observe that $\det(P) = \det(P_1 P_2)$; see, *e.g.*, [102, Lemma 2.1.]. Therefore,

$$-\log\det(P) = -\log\det(P_1) - \log\det(P_2)$$

which is a convex function on $\mathcal{S}_+^n \times \mathcal{S}_+^n$. Thus, Problem 3.2 can be stated as follows.

$$\begin{aligned} &\underset{P_1,P_2,S_1,S_2,J,\tau,K}{\text{minimize}} && -\log\det(P_1) - \log\det(P_2) \\ &\text{subject to} && S_1, S_2 \in \mathcal{D}_+^p, P_1, P_2 \in \mathcal{S}_+^n, \tau > 0 \\ & && (3.17), (3.18). \end{aligned} \tag{3.40}$$

On the other hand, as pointed out earlier, the adoption of the latter criterion could lead to a set $\mathcal{E}(P)$ excessively stretched along some directions. To overcome this problem, as already done in Chapter 2, instead of minimizing the volume of $\mathcal{E}(P)$, one can minimize trace(P^{-1}). However, since this criterion is in general non-convex in the decision variables P_1, P_2, its exploitation in a numerical scheme is not straightforward. To overcome this drawback, we introduce a further variable $N \in \mathcal{S}_+^{2n}$, subject to the following linear constraint

$$\begin{bmatrix} N & \mathrm{I} \\ \bullet & P \end{bmatrix} \geq \mathbf{0}$$

which, by Schur complement, is equivalent to $P^{-1} \leq N$. Therefore, the minimization of $\mathrm{trace}(P^{-1})$ can be implicitly performed by minimizing $\mathrm{trace}(N)$, which is a convex (in fact linear), function. By pursuing this approach, Problem 3.2 turns in

$$
\begin{aligned}
\underset{P_1,P_2,S_1,S_2,J,\tau,N,K}{\text{minimize}} \quad & \mathrm{trace}(N) \\
\text{subject to} \quad & S_1, S_2 \in \mathcal{D}_+^p, P_1, P_2 \in \mathcal{S}_+^n, N \in \mathcal{S}_+^{2n}, \tau > 0 \\
& \begin{bmatrix} N & \mathrm{I} \\ \bullet & P \end{bmatrix} \geq \mathbf{0} \\
& (3.17), (3.18).
\end{aligned}
\tag{3.41}
$$

Another alternative solution, consists of minimizing the set $\mathcal{E}(P)$ along certain directions of interest. In particular, let $v_1, v_2, \dots, v_p \in \mathbb{R}^{2n}$ be some given vectors, and let $\theta_1, \theta_2, \dots, \theta_p$, positive scalars. Consider the following constraints

$$
v_i^{\mathsf{T}} P v_i \geq \theta_i \qquad i = 1, 2, \dots, s. \tag{3.42}
$$

By maximizing the scalars θ_i, the set $\mathcal{E}(P)$ shrinks along the directions v_i. In this case, Problem 3.2 can be stated as follows

$$
\begin{aligned}
\underset{P_1,P_2,S_1,S_2,J,\tau,K,\theta_1,\theta_2,\dots,\theta_s}{\text{minimize}} \quad & -\sum_{i=1}^{s} \theta_i \gamma_i \\
\text{subject to} \quad & S_1, S_2 \in \mathcal{D}_+^p, P_1, P_2 \in \mathcal{S}_+^n, \tau > 0 \\
& (3.17), (3.18), (3.42).
\end{aligned}
\tag{3.43}
$$

where $\gamma_i > 0$ are the weights of the different objectives.

Remark 3.1. Notice that the results derived in this chapter aim at characterizing the whole control system state, *i.e.*, (x, x_c). However, the x_c component of the state is somehow artificial, and one may be in general more interested in the behavior of the plant state. Nonetheless, the application of the presented results allows to draw some conclusions on the plant state x. Specifically, notice that UGAS of the set $\mathcal{A} = \mathcal{E}(P)$ entails global attractivity of the set

$$
\mathcal{A}_x \coloneqq \{x \in \mathbb{R}^n \colon x^{\mathsf{T}} P_1 x \leq 1\} \times \mathbb{R}^n.
$$

To see this, observe that[2]

$$
P = \begin{bmatrix} P_1 + P_2 & -P_2 \\ \bullet & P_2 \end{bmatrix} \geq \begin{bmatrix} P_1 & \mathbf{0} \\ \bullet & \mathbf{0} \end{bmatrix}
$$

and the latter implies that $\mathcal{A} \subset \mathcal{A}_x$, yielding the attractivity of $\mathcal{A}_x$. However, uniform stability of $\mathcal{A}_x$ cannot be established being such a set, in general, not strongly forward invariant for (3.8)[3]. Building on this observation, with the aim steering the plant state as

[2]Since $P_2 > \mathbf{0}$, this inequality readily follows from the application of the Schur complement lemma for nonstrict inequalities; see [15].

[3]Notice that, by definition, uniform stability of a given set implies its strongly forward invariance.

much as possible to the origin, one may suitably modify both the determinant and the trace criteria illustrated above by only focusing on the matrix P_1.

Numerical Issues in the Solution to Problem 3.2

As pointed out earlier, (3.37) is still nonlinear in the decision variables due to the terms P_1BK (and its symmetric), and τP_1. While the latter nonlinearity can be easily managed via a grid search, the first is hardly tractable. To be best author knowledge, there are no strategies in the literature to perform the design of a linear observer-based controller through the solution to linear matrix inequalities. On the other hand, Fact 3.1 ensures the feasibility of the conditions given by Proposition 3.3 under the Hurwitzness of the matrices $A + BK$. Hence, by assuming that the controller gain K is a given stabilizing gain, (3.17) can be used in a convex setup to design the observer gain without leading to any drawback in terms of feasibility of the resulting optimization problems.

Remark 3.2. The selection of the controller gain K somehow constraints the feasibility set of the above optimization problems. Indeed, once K is given, to ensure the feasibility of (3.18), $\tau \in (0, 2 \min \mathcal{R}(A + BK))$, where $\mathcal{R}(A + BK) \coloneqq \{|\Re(\lambda)| \colon \lambda \in \operatorname{spec}(A + BK)\}$.

Remark 3.3. As pointed out in Chapter 2, approaching the optimal solutions to Problem 3.2 may lead to solutions characterized by a large gain L. Such a situation needs to be avoided to envision the physical construction of the proposed controller. As already discussed in Chapter 2, a general way to overcome this drawback consists of adding suitable constraints on the eigenvalues of the matrix $A - LC$.

Notice that, as long as the considered objective function is convex, whenever the scalar τ is fixed and a choice is considered for the gain K, the above proposed optimization problems are genuine convex optimization problems over LMIs. On the other hand, as for the matter of the optimization problems presented in Chapter 2, the positive scalar τ can be treated as a tuning parameter, or being selected via an grid search. Moreover, Fact 3.1 provides a valuable tool to characterize the interval of values for τ ensuring the feasibility of (3.17) and (3.18). Based on this idea, consider the following algorithm, that by performing a search on certain interval for τ (wherein the feasibility of the considered optimization problem is ensured), provides a possible solution to determine a (sub)optimal solution to Problem 3.2.

Algorithm 3.1 Observer-based controller design

Input: Matrices A, B, K, C, scalars $\Delta > 0$, a convex criterion $\mathcal{M}_s$, and a tolerance $\rho > 0$.

Initialization: Let $\mathcal{R}(A+BK) := \{|\Re(\lambda)| \colon \lambda \in \operatorname{spec}(A+BK)\}$, select $\tau = 2 \times 0.99 \min \mathcal{R}(A+BK)$,

Iteration

Step 1:

Solve the following convex optimization problem over LMIs

$$\begin{aligned} &\underset{S_1,S_2,P_1,P_2,J}{\text{minimize}} && \mathcal{M}_s(P_1,P_2) \\ &\text{s.t.} && S_1, S_2 \in \mathcal{D}_+^m, P \in \mathcal{S}_+^n \\ &&& \begin{bmatrix} \mathrm{He}(P_1(A+BK)) + \tau P_1 & -P_1 BK & -C^\mathsf{T} S_2 \\ \bullet & \mathrm{He}(P_2 A - JC) + \tau P_2 & -J \\ \bullet & \bullet & -2S_2 - S_1 \end{bmatrix} < \mathbf{0} \\ &&& \operatorname{trace}(S_1)\Delta^2 - \tau \leq 0 \end{aligned}$$

Pick the sub-optimal solution $(\bar{P}_1, \bar{P}_2, \bar{J})$ and store the obtained solution:

$$\mathcal{M}_{s\star}^{(k)} \leftarrow \mathcal{M}_s(\operatorname{diag}\{\bar{P}_1, \bar{P}_2\}), P_{1\star}^{(k)} \leftarrow \bar{P}_1, P_{2\star}^{(k)} \leftarrow \bar{P}_2, J_\star^{(k)} \leftarrow \bar{J}.$$

$k \leftarrow k+1$

Step 2:

Decrease τ of ρ, *i.e.*, $\tau \leftarrow \tau - \rho$

Until $\tau > 0$.

Step 3: $k_{\max} \leftarrow k$, select $k^\star = \underset{k \in \{1,2,k_{\max}\}}{\operatorname{argmin}} \{\mathcal{M}_{s\star}^{(k)}\}$

Output: $P_1 = P_{1\star}^{(k^\star)}, P_2 = P_{2\star}^{(k^\star)}, L = P_2^{-1} J_\star^{(k^\star)}$

It is worthwhile to remark that, the above algorithm, due to the proposed initialization arising from the proof of Fact 3.1, always terminates with a suboptimal solution to Problem 3.2.

Numerical Example

Example 3.1. Consider the system derived from [49], already considered in Example 2.4, that is defined by the following data:

$$A = \begin{bmatrix} 0 & 1 \\ 0.5 & 0.5 \end{bmatrix}, B = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, K = \begin{bmatrix} -0.3491 & -0.7022 \end{bmatrix}.$$

Assume that the plant state is not fully accessible, and that only the second component is measured via a uniform quantized sensor with $\Delta = 1$, *i.e.*, $y_m = \mathrm{q}\left(\begin{bmatrix} 0 & 1 \end{bmatrix} x\right)$. We want to solve Problem 3.3 via Algorithm 3.1, by using the trace criterion as presented in (3.60). In particular, let $N \in \mathcal{S}_+^{2n}$, by requiring that

$$\begin{bmatrix} N & \mathrm{I} \\ \bullet & \mathrm{diag}(P_1, P_2) \end{bmatrix} \geq \mathbf{0}$$

as convex objective to minimize, we consider $\mathrm{trace}(N)$. By considering $\rho = 0.01$, Algorithm 3.1 yields

$$\begin{aligned} P_1 &= \begin{bmatrix} 1.05716438 & -1.05651233 \\ -1.05651233 & 1.05766293 \end{bmatrix} \\ P_2 &= \begin{bmatrix} 8.11663119 & -8.11833684 \\ -8.11833684 & 8.12644359 \end{bmatrix} \\ L &= \begin{bmatrix} 9.3755 \\ 9.3735 \end{bmatrix} \\ &\mathrm{trace}(\mathrm{diag}\{P_1^{-1}, P_2^{-1}\}) \approx 1422. \end{aligned}$$

Now, with the aim of steering the plant state as much as possible to the origin, we want to solve Problem 3.3 by considering a trace criterion based only on P_1. In particular, as mentioned earlier, let $N \in \mathcal{S}_+^n$, by requiring that

$$\begin{bmatrix} N & \mathrm{I} \\ \bullet & P_1 \end{bmatrix} \geq \mathbf{0}$$

as convex objective to minimize, we consider $\mathrm{trace}(N)$. Since in this case the matrix P_2 in not accounted by the size criteria and its inversion is needed to derive the gain L, to avoid numerical problems, in the solution of the considered optimization problem, we consider additional constraints on the matrix P_2 aimed at ensuring a strictly positive lower bound on

$\lambda_{\min}(P_2)$. That said, by selecting $\rho = 0.01$, Algorithm 3.1 yields

$$P_1 = \begin{bmatrix} 2.02467621 & -2.02376235 \\ -2.02376235 & 2.0252034 \end{bmatrix}$$
$$P_2 = \begin{bmatrix} 6.2456427 & -36.2469778 \\ -36.2469778 & 36.2485599 \end{bmatrix}$$
$$L = \begin{bmatrix} 200.40 \\ 200.36 \end{bmatrix}$$
$$\mathrm{trace}(P_1^{-1}) \approx 849.54.$$

Figure 3.1 shows the evolution of the plant state obtained by considering the two different design. In both simulations, the closed-loop system is initialized as $(x_0, \hat{x}_0) = (-6, 0, 0, 0)$. Simulations show that, although the controller gain K is the same in both cases, the second design allows to steer the plant state closer to the origin.

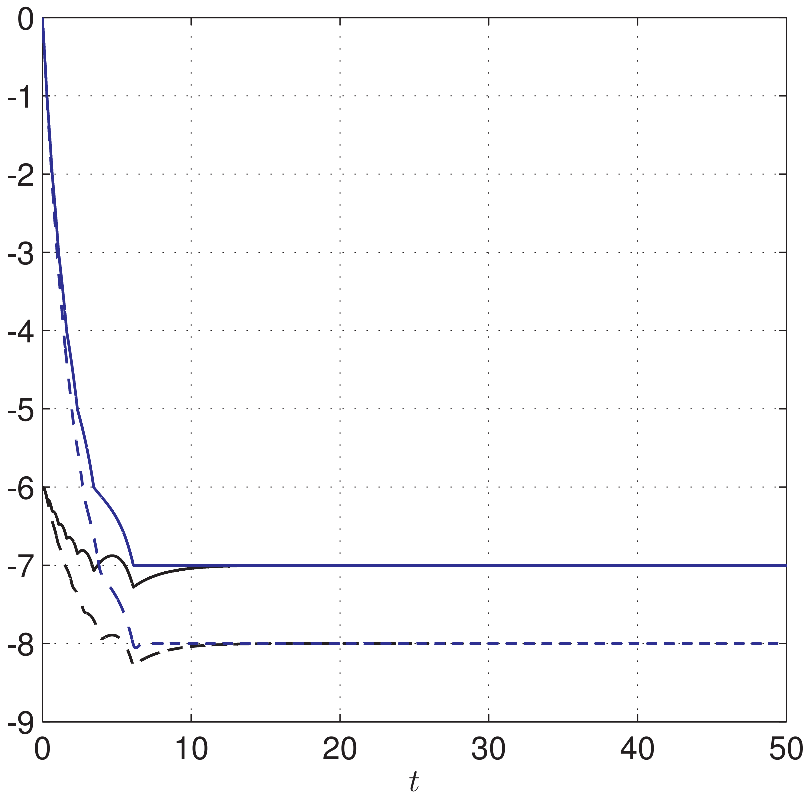


Figure 3.1: Plant state evolution: First design (x_1 dashed-black, x_2 dashed-blue), second design (x_1 solid-black x_2 solid-blue). The solutions are obtained by integrating the closed-loop model via an Euler method with time step 10^{-4}.

3.2.4 Full Dynamic Controller Design

The aim of this subsection is to provide another design strategy for the controller in (3.2), which avoids the above illustrated issues preventing from designing the whole controller via the solution to a convex optimization problem. The approach followed in this section revolves around the congruence transformations and the change of variables presented in [25]. In particular, as a first step, we give an equivalent condition to (3.10), which is linear in the (new) decision variables, whenever τ and S_2 are fixed. Such a result, whose proof is a slight variation of the one of [25, Theorem 4.3], is given next.

Proposition 3.4. *For each τ, $S_1, S_2 \in \mathcal{D}_+^p$, there exist $X, Y \in \mathcal{S}_+^n$, $K \in \mathbb{R}^{n\times n}$, $L \in \mathbb{R}^{n\times p}$, $M \in \mathbb{R}^{m\times n}$, $N \in \mathbb{R}^{m\times p}$ such that*

$$\begin{bmatrix} \mathrm{He}(H_1) + \tau H_2 & H_3 \\ \bullet & -S_1 - 2S_2 \end{bmatrix} < \mathbf{0} \tag{3.44}$$

$$H_2 > \mathbf{0} \tag{3.45}$$

where

$$H_1 = \begin{bmatrix} AY + BM & A + BNC \\ K & XA + LC \end{bmatrix}, H_2 = \begin{bmatrix} Y & \mathrm{I} \\ \bullet & X \end{bmatrix}, H_3 = \begin{bmatrix} BN - YC^{\mathsf{T}} S_2 \\ L - C^{\mathsf{T}} S_2 \end{bmatrix}$$

if and only if, for any nonsingular matrices $U, V \in \mathbb{R}^{n\times n}$ such that $UV^{\mathsf{T}} = \mathrm{I} - XY$, $\hat{X} = U^{\mathsf{T}}(X - Y^{-1})^{-1}U$, the matrices

$$\begin{aligned} D_c &= N \\ C_c &= (M - NCY)V^{-\mathsf{T}} \\ B_c &= U^{-1}(L - XBN) \\ A_c &= U^{-1}\left(K - XAY - XBM - UB_cCY\right)V^{-\mathsf{T}} \end{aligned} \tag{3.46}$$

and

$$P = \begin{bmatrix} X & U \\ \bullet & \hat{X} \end{bmatrix} \tag{3.47}$$

satisfy (3.10).

Proof. (Necessity) We want to prove that (3.10), implies (3.44) and (3.45).
Let $P \in \mathcal{S}_+^{2n}$, and A_c, B_c, C_c, D_c matrices of adequate dimensions, such that (3.10) is verified. Let us denote

$$P = \begin{bmatrix} X & U \\ \bullet & \hat{X} \end{bmatrix}, P^{-1} = \begin{bmatrix} Y & V \\ \bullet & \hat{Y} \end{bmatrix}. \tag{3.48}$$

Thus, the following relations hold

$$XY + UV^{\mathsf{T}} = \mathrm{I} \tag{3.49}$$
$$XV + U\hat{Y} = \mathbf{0} \tag{3.50}$$
$$U^{\mathsf{T}}V + \hat{X}\hat{Y} = \mathrm{I} \tag{3.51}$$
$$U^{\mathsf{T}}Y + \hat{X}V^{\mathsf{T}} = \mathbf{0} \tag{3.52}$$

in particular, $UV^{\mathsf{T}} = \mathrm{I} - XY$, and $\hat{X} = U^{\mathsf{T}}(X - Y^{-1})^{-1}U$. Define

$$\mathbb{J} = \begin{bmatrix} Y & V \\ \mathrm{I} & \mathbf{0} \end{bmatrix}$$

and observe that, as shown in [25, Lemma 4.2], U, V can be assumed, without loss of generality, nonsingular. This latter assumption assures nonsingularity of $\mathbb{J}$ (see Lemma A.2). Pre-and-post multiplying the left-hand side of (3.10) respectively by $\mathrm{diag}(\mathbb{J}, \mathrm{I})$ and $\mathrm{diag}(\mathbb{J}^{\mathsf{T}}, \mathrm{I})$, from the satisfaction of (3.10) it follows

$$\begin{bmatrix} \mathrm{He}(\mathbb{J}P\widetilde{A}\mathbb{J}^{\mathsf{T}}) + \tau\mathbb{J}P\mathbb{J}^{\mathsf{T}} & \mathbb{J}P\widetilde{B} - \mathbb{J}\widetilde{C}^{\mathsf{T}}S_2 \\ \bullet & -2S_2 - S_1 \end{bmatrix} < \mathbf{0} \tag{3.53}$$

where

$$\mathbb{J}P = \begin{bmatrix} \mathrm{I} & \mathbf{0} \\ X & U \end{bmatrix} \tag{3.54}$$

$$\mathbb{J}P\mathbb{J}^{\mathsf{T}} = \begin{bmatrix} Y & \mathrm{I} \\ I & X \end{bmatrix} \tag{3.55}$$

$$\mathbb{J}P\widetilde{A}\mathbb{J}^{\mathsf{T}} = \begin{bmatrix} (A + BD_cC)Y + BC_cV^{\mathsf{T}} & A + BD_cC \\ X(A + BD_cC)Y + XBC_cV^{\mathsf{T}} + UB_cCY + UA_cV^{\mathsf{T}} & X(A + BD_cC) + UB_cC \end{bmatrix} \tag{3.56}$$

$$\mathbb{J}P\widetilde{B} = \begin{bmatrix} BD_C \\ XBD_c + UB_c \end{bmatrix} \tag{3.57}$$

$$\mathbb{J}\widetilde{C}^{\mathsf{T}} = \begin{bmatrix} YC^{\mathsf{T}} \\ C^{\mathsf{T}} \end{bmatrix} \tag{3.58}$$

then $\mathbb{J}P\mathbb{J}^{\mathsf{T}} = H_2 > \mathbf{0}$ yielding (3.45). Now, let us consider the following change of variables given in [25]

$$\begin{bmatrix} K & L \\ M & N \end{bmatrix} = \begin{bmatrix} XAY & \mathbf{0} \\ \bullet & \mathbf{0} \end{bmatrix} + \begin{bmatrix} U & XB \\ \mathbf{0} & \mathrm{I} \end{bmatrix} \begin{bmatrix} A_c & B_c \\ C_c & D_c \end{bmatrix} \begin{bmatrix} V^{\mathsf{T}} & \mathbf{0} \\ CY & \mathrm{I} \end{bmatrix}. \tag{3.59}$$

By straightforward calculations, it turns out that

$$\begin{aligned}
&\mathbb{J}P\widetilde{A}\mathbb{J}^{\mathsf{T}} = H_1 \\
&\mathbb{J}P\mathbb{J}^{\mathsf{T}} = H_2 \\
&\mathbb{J}P\widetilde{B} - \mathbb{J}\widetilde{C}^{\mathsf{T}} S_2 = H_3
\end{aligned}$$

Therefore, it follows that the satisfaction of (3.10) implies the satisfaction of (3.44).

(Sufficiency) Let $X, Y \in \mathcal{S}_+^n$, $K \in \mathbb{R}^{n\times n}$, $L \in \mathbb{R}^{n\times p}$, $M \in \mathbb{R}^{m\times n}$, $N \in \mathbb{R}^{m\times p}$ such that (3.44) and (3.45) are verified. Then, it is always possible to determine U, V nonsingular such that $\mathrm{I} - XY = UV^{\mathsf{T}}$, and also such that

$$\mathbb{J} = \begin{bmatrix} Y & V \\ \mathrm{I} & \mathbf{0} \end{bmatrix}$$

is nonsingular (see Lemma A.2). Now, from (3.45), as shown in the necessity part, since $\mathbb{J}$ is nonsingular, it follows that

$$P = \mathbb{J}^{-1} H_2 \mathbb{J}^{-\mathsf{T}} = \begin{bmatrix} X & U \\ \bullet & \hat{X} \end{bmatrix} > \mathbf{0}.$$

To conclude, it suffices to observe that, due to U nonsingular, the change of variables in (3.59) is invertible. In particular, by inversion of the relation given in (3.59), one gets the relations in (3.46). Now, recall that (3.44) was derived in the necessity part from (3.10) by the change of variables in (3.59), and a congruence transformation involving the matrix $\mathbb{J}$. Hence, the satisfaction of (3.44) implies the satisfaction of (3.10), with P given in (3.47), and A_c, B_c, C_c, D_c given in (3.46), and this concludes the proof. ■

Optimization Issues

As already mentioned, in solving Problem 3.1, the main objective consists of designing the controller (3.2) to ensure that the closed-loop solutions converge sufficiently close to the origin. To this end, building on the conditions provided by Proposition 3.1, one can consider the following optimization problem.

Problem 3.3 (Stabilization)**.** Let A, B, C be matrices of adequate dimensions.
Determine A_c, B_c, C_c, D_c, and $P \in \mathcal{S}_+^{2n}$, such that $\mathcal{E}(P)$ is UGAS for system (3.8), and it is minimized with respect to some criterion.

As already illustrated in the previous chapters, the solution to the above optimization problem can be carried out by embedding the conditions provided by Proposition 3.4 into a suitable optimization scheme. To this end, an adequate measure of the set $\mathcal{E}(P)$ needs to be selected. Differently from the previous chapters, in this setting, since P is nondiagonal, the adoption of a criterion based on the determinant of P would give rise, in general, to a non-convex criterion. For this reason, as a criterion to be minimized, in this chapter, we consider $\mathrm{trace}(P^{-1})$. In particular, notice that from Proposition 3.4, it follows that

$$P^{-1} = \begin{bmatrix} Y & V \\ \bullet & \widehat{Y} \end{bmatrix}$$

where $V = (\mathrm{I} - XY)^T U^{-\mathsf{T}}$, and $\widehat{Y} = -U^{-1}XV$, which in turn yields $\widehat{Y} = -U^{-1}X(\mathrm{I} - XY)^T U^{-\mathsf{T}}$. At this stage, observe that modulo the nonsingularity requirement, U can be arbitrarily chosen without any influence on the feasibility of the conditions given by Proposition 3.4. Therefore, building on this degree of freedom, U can be selected with the aim of determining a convex criterion in the decision variables. In particular, selecting $U = X$ yields

$$\mathrm{trace}(P^{-1}) = 2\,\mathrm{trace}(Y) - \mathrm{trace}(X^{-1})$$

and the latter expression points out that $\mathrm{trace}(P^{-1})$ can be implicitly minimized by simultaneously minimizing $\mathrm{trace}(Y)$, and $-\mathrm{trace}(X^{-1})$. On the other hand, the minimization of $-\mathrm{trace}(X^{-1})$, being $X \in \mathcal{S}^n_+$, can be indirectly obtained by minimizing $\mathrm{trace}(X)$. By pursuing this strategy, Problem 3.3 specializes in

$$\begin{array}{ll} \underset{X,Y,L,K,M,N,\tau,S_1,S_2}{\text{minimize}} & \mathrm{trace}(X+Y) \\ \text{subject to} & S_1, S_2 \in \mathcal{D}^p_+, X, Y \in \mathcal{S}^n_+, \tau > 0 \\ & (3.44), (3.45), (3.11). \end{array} \tag{3.60}$$

Remark 3.4. Another convex criterion can be worked out by following a similar strategy to the one in [53]. However, establishing which of the two strategies always provides the best result is difficult.

Numerical Issues in the Solution to (3.60)

Notice that, due to the terms τX, τY, and $YC^{\mathsf{T}}S_2$ (and its symmetric) (3.44) is nonlinear in the decision variables. Therefore, from a numerical standpoint, the solution to (3.60) may lead to NP-hard problems. Nevertheless, whenever τ and S_2 are fixed (3.44) turns into a genuine LMI. As already discussed throughout this dissertation, the nonlinear terms τX and τY can be easily managed by performing a grid search for τ over a certain interval $(0, \tau_{\max})$. Instead, the selection of S_2 could be more complicated. However, notice that $S_2 \in \mathcal{D}^p_+$, hence at least for $p \leq 2$, a grid search for S_2 represents a viable strategy to determine a solution to (3.60). Differently from other cases treated in this dissertation, the derivation of linear sufficient conditions to (3.44) appears hard due to the increased complexity of (3.44) with respect to the simpler conditions presented in Chapter 2. On the other hand, another viable strategy arises from the combined exploitation of Proposition 3.2 and Proposition 3.4. Such a strategy is schematized as follows:

- as a first step, select some stabilizing controller for the triplet (A, B, C); this is always possible due to Assumption 1.2.

- once the controller is known, by fixing τ as prescribed in the proof of Proposition 3.2, (3.44) becomes a genuine LMI in the remaining variables, whose feasible set is nonempty. Therefore, S_2 can be selected to ensure the feasibility of (3.44)-(3.45) under the choices considered in the first step for τ and for the controller.

- once S_2 is known, by preforming a grid search for τ, one can derive a suboptimal solution to (3.60) by solving a finite numbers of LMI optimization problems.

Remark that, since Proposition 3.4 provides an equivalent condition to (3.44) and the choice of S_2, and τ ensures the feasibility of (3.44)-(3.45) for the controller chosen to start the procedure, it follows that throughout the third stage of the above procedure, the feasible set of (3.60) is nonempty.

The above idea is concretely adopted to develop the following algorithm. Such an algorithm, analogously to Algorithm 2.2, performs an improved search for τ.

Algorithm 3.2 Controller design

Input: Matrices A, B, C, scalar $\Delta > 0$, and a tolerance $\rho > 0$.
Initialization: Select $A_c^{(0)}, B_c^{(0)}, C_c^{(0)}, D_c^{(0)}$, such that

$$\widetilde{A}^{(0)} = \begin{bmatrix} A + BD_c^{(0)}C & BC_c^{(0)} \\ B_c^{(0)}C & A_c^{(0)} \end{bmatrix}$$

is Hurwitz. Let $\mathcal{R}(\widetilde{A}^{(0)}) := \{|\Re(\lambda)| \colon \lambda \in \operatorname{spec}(\widetilde{A}^{(0)})\}$. Set for the next step

$$\overline{\tau} = 2 \times 0.99 \min \mathcal{R}(\widetilde{A}^{(0)}) \quad \widetilde{B}^{(0)} = \begin{bmatrix} BD_c \\ B_c \end{bmatrix}$$

Step 1:
Determine a feasible solution to the following LMI problem

$$\begin{aligned} & S_1, S_2 \in \mathcal{D}_+^p, P \in \mathcal{S}_+^{2n} \\ & \begin{bmatrix} \mathrm{He}(P\widetilde{A}^{(0)}) + \overline{\tau}P & P\widetilde{B}^{(0)} - \widetilde{C}^\mathsf{T} S_2 \\ \bullet & -2S_2 - S_1 \end{bmatrix} < \mathbf{0} \\ & \operatorname{trace}(S_1)\Delta^2 - \overline{\tau} \leq 0 \end{aligned}$$

Set $\overline{S}_2 = S_2$ for the next step. Select a grid of positive values $\mathcal{G}_\tau$ such that $\bar{\tau} = \max \mathcal{G}_\tau$
Iteration
Step 2:
Solve the following LMI optimization problem selecting τ over $\mathcal{G}_\tau$

$$\begin{aligned} \underset{X,Y,L,K,M,N,S_1}{\text{minimize}} \quad & \operatorname{trace}(X+Y) \\ \text{subject to} \quad & S_1 \in \mathcal{D}_+^p, X, Y \in \mathcal{S}_+^n, \operatorname{trace}(S_1)\Delta^2 - \tau \leq 0, H_2 > \mathbf{0} \\ & \begin{bmatrix} \mathrm{He}\left(\left[\begin{smallmatrix} AY+BM & A+BNC \\ K & XA+LC \end{smallmatrix}\right]\right) + \tau \left[\begin{smallmatrix} Y & \mathrm{I} \\ \bullet & X \end{smallmatrix}\right] & \left[\begin{smallmatrix} BN - YC^\mathsf{T}\overline{S}_2 \\ L - C^\mathsf{T}\overline{S}_2 \end{smallmatrix}\right] \\ \bullet & -S_1 - 2\overline{S}_2 \end{bmatrix} < \mathbf{0} \end{aligned}$$

Pick the suboptimal solution to the above optimization problem

$$\mathcal{X}^\star \leftarrow (\tau^\star, X^\star, Y^\star, L^\star, K^\star, M^\star, N^\star, S_1^\star).$$

Set $U = X^\star$, $V = (\mathrm{I} - X^\star Y^\star)^\mathsf{T} U^{-\mathsf{T}}$, and determine the controller parameters from $\mathcal{X}^\star$ via (3.46), and P via (3.47).
Determine the closed-loop matrix $\widetilde{A}$, and set $\overline{\tau} = 2 \times 0.99 \min \mathcal{R}(\widetilde{A})$. Build a grid of positive values $\mathcal{G}_\tau$ such that $\bar{\tau} = \max \mathcal{G}_\tau$, and $\tau^\star \in \mathcal{G}_\tau$, (notice that necessarily $\tau^\star \leq \bar{\tau}$. Including $\tau^\star$ in $\mathcal{G}_\tau$ ensures the feasibility at each step).

Until $\operatorname{trace}(X+Y)$ does not decrease below ρ over three consecutive steps.

Output: (A_c, B_c, C_c, D_c, P)

Obviously the above algorithm may converge to different solutions depending on the controller chosen throughout the initialization stage.

Remark 3.5. The initializing controller required to start the above algorithm can be designed via standard linear techniques as LQG control design. Another solution consists of selecting as initializing controller, the observer-based controller built through the apparatus proposed in Section 3.2.3.

Remark 3.6. Notice that the design stage, to be numerically tractable, introduces some conservatism in the determination of the set $\mathcal{A} = \mathcal{E}(P)$. Essentially, this additional conservatism depends on the fact that the selection of S_2 can dramatically affect the achievable suboptimal solutions. Therefore, as a second step, one can envision a further analysis stage to obtain a tighter estimation of the real behavior of the closed-loop system. Such a stage can be performed by embedding the conditions provided by Proposition 3.1 in a suitable optimization problem. In particular, once the controller parameters are known, and some selection for τ is considered, relation (3.10) turns into a genuine LMI. Thus, provided that a convex criterion to measure the set $\mathcal{A}$ is chosen, a potentially tighter set $\mathcal{A}$ can be determined by solving a finite number of convex optimization problems over LMIs.

As pointed out earlier in this dissertation, to attain the optimal solutions to Problem 3.3 the controller parameters could even blowup (see [110] for a formal treatment of these issues in the case of $\mathcal{H}_\infty$ state feedback control). Obviously, such a situation needs to be avoided to envision the physical construction of the proposed controller. To overcome this problem, one may consider further constraints aimed at placing the eigenvalues of the matrix $\widetilde{A}$ in certain sectors of the complex plane. Via the apparatus proposed in [25], such constraints can be easily integrated in the solution to (3.60) by means of additional LMIs in the decision variables. Some classical constraints, along with sufficient conditions (linear in the decision variables of (3.60)), are given below.

- Disk centered at $(-q, 0)$ with radius r

$$\begin{bmatrix} -rH_2 & qH_2 + H_1 \\ \bullet & -rH_2 \end{bmatrix} < \mathbf{0} \tag{3.61}$$

- Open-half plane $\Re(z) < -\alpha$

$$2\alpha H_2 + \mathrm{He}(H_1) < \mathbf{0} \tag{3.62}$$

- Open-half $\Re(z) > -\alpha$

$$-2\alpha H_2 - \mathrm{He}(H_1) < \mathbf{0} \tag{3.63}$$

- Conic sector with apex at the origin and inner angle 2θ

$$\mathrm{He}\left(\begin{bmatrix} \sin(\theta)H_1 & \cos(\theta)H_1 \\ -\cos(\theta)H_1 & \sin(\theta)H_1 \end{bmatrix}\right) < \mathbf{0}. \tag{3.64}$$

Remark 3.7. The selection of the more convenient pole placement constraint needs to be tailored to the considered case. Also, notice that including the above constraints may impact on the feasibility of the resulting optimization problem, and dramatically affect the value of the achievable suboptimal solution. In other words, the addition of pole placement constraints somehow reshapes the feasible set of the considered optimization problem in a way that appears unclear. However, in general, one can reasonably assume that, as long as the pole placement constraint is not excessively severe, at least the feasibility of the resulting optimization problem should be preserved. This aspect was already discussed in Chapter 2 for the matter of the actuator quantization case. In that setting, we shown a possible strategy to limit the effect induced by pole placement constraints for the considered optimization problems. On the other hand, pursuing this approach in the case under study in this section appears nontrivial.

Numerical Examples

Example 3.2. Consider the balancing pointer system derived from [69] that is defined by the following data:

$$A = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, B = \begin{bmatrix} 0 \\ -1 \end{bmatrix}.$$

Assume that the plant state is not fully accessible, and that only the first component is measured via a uniform quantized sensor with $\Delta = 0.5$, *i.e.*, $y_m = \mathrm{q}\left(\begin{bmatrix} 1 & 0 \end{bmatrix} x\right)$. We want to solve Problem 3.3 by performing a simultaneous grid search for the positive scalars τ and S_2. Moreover to avoid the occurrence of fast dynamics or/and high gains in the designed controller, we consider, for the matrix $\widetilde{A}$ the pole placement constraint in (3.63) with $\alpha = 10$. The latter constraint provides an indication on how to choose the upper bound defining the grid of values for τ inspected throughout the design stage. Indeed, since (3.63) with $\alpha = 10$ implies that for each $\lambda \in \mathrm{spec}(\widetilde{A})$, $\Re(\lambda) > -10$, to ensure the feasibility of the considered optimization problem, it has to be $\tau < 20$.

Concerning the choice of the grid of values for S_2, bearing in mind that, as shown in the proof of Proposition 2.2, selecting $S_2 = 0$ ensures the feasibility of (3.10) and (3.11) for some $P \in \mathcal{S}_+^{2n}$, $S_1 \in \mathcal{D}_+^p$, and $\tau > 0$, (at least when no additional constraints on the eigenvalues of $\widetilde{A}$ are considered), it follows that S_2 can be selected small enough and then increased up to a certain value to meet the desired optimization specifications. However, the selection of S_2 strongly depends on the considered cases, and a systematic algorithm for its selection appears complicated. Hence, a certain tuning stage for this variable needs to be considered.

In this case, keeping in mind the constraint on τ, we let τ vary over a grid of 50 points from 0.1 up to 20×0.99. For S_2, we still consider a grid of 50 points from 0.001 up to 0.1. Figure 3.2 depicts the evolution of the optimal value of $\mathrm{trace}(X + Y)$ obtained by solving (3.60) over the grid chosen for τ and S_2, versus τ and S_2. The figure empathizes that the value of the suboptimal solution strongly depends on the values chosen for τ and S_2, and that due to the further constraint ensuring the desired pole placement, the largest value of τ ensuring the feasibility of (3.60) is smaller than the upper bound considered in the related grid. Specifically, the most convenient values selected for τ and S_2 are $\tau = 0.904, S_2 = 0.0232$, that give

$$
\begin{aligned}
A_c &= \begin{bmatrix} -3.232 & 1.379 \\ -9.587 & -11.56 \end{bmatrix} \\
B_c &= \begin{bmatrix} -3.334 \\ 5.174 \end{bmatrix} \\
C_c &= \begin{bmatrix} -7.433 & -11.86 \end{bmatrix} \\
D_c &= 8.729 \\
P &= \begin{bmatrix} 1.741 & -0.9672 & 1.741 & -0.9672 \\ -0.9672 & 1.082 & -0.9672 & 1.082 \\ 1.741 & -0.9672 & 2628 & 1692 \\ -0.9672 & 1.082 & 1692 & 1093 \end{bmatrix}.
\end{aligned}
$$

For such a solution, one has $\operatorname{trace}(X+Y) \approx 10.5549$.

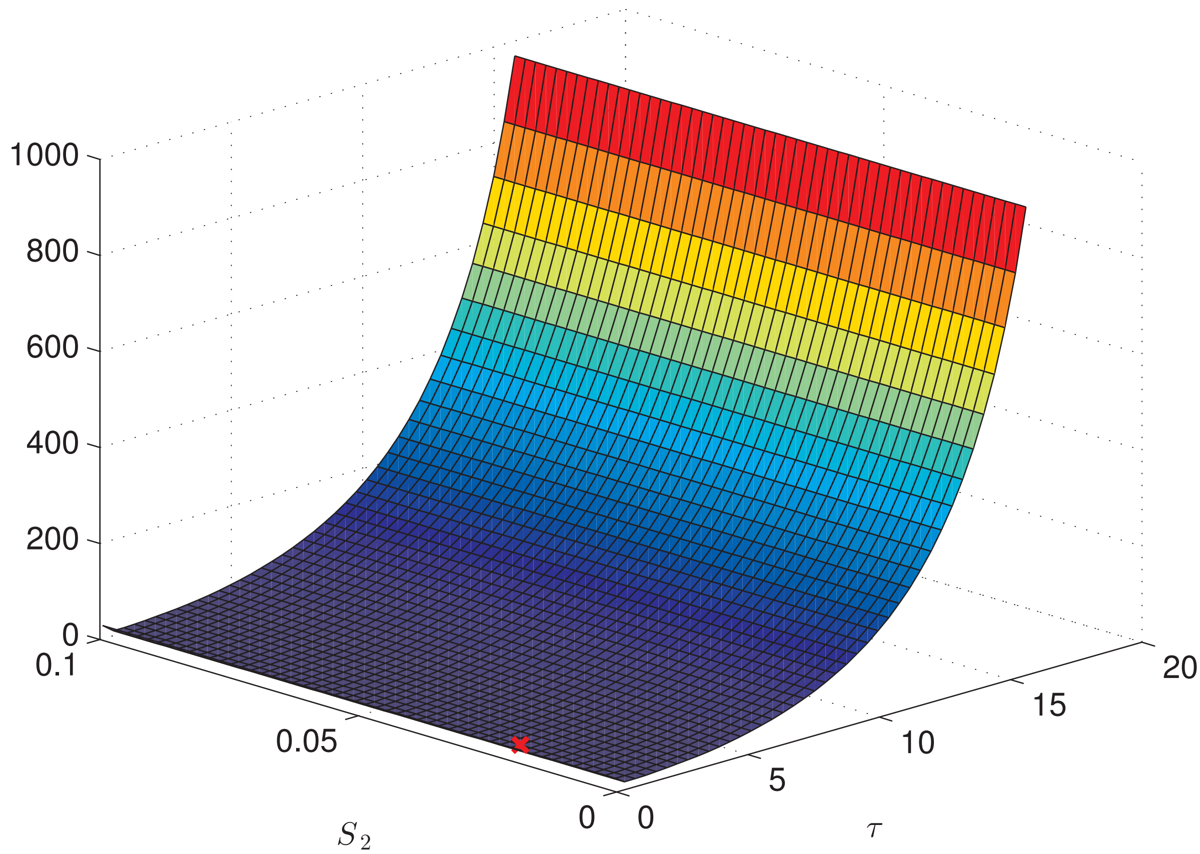


Figure 3.2: The optimal value of $\operatorname{trace}(X+Y)$ obtained by solving (3.60) over the grid chosen for τ and S_2, vs τ and S_2. The red cross indicates the suboptimal solution to (3.60).

To foster the use of the pole placement constraints mentioned above, we want to show that the solution to (3.60), whenever no pole placement constraints are considered, can lead to controllers in practice unattainable. To this end, still consider $\tau = 0.904, S_2 = 0.0232$, the solution to (3.60) gives

$$
\begin{aligned}
A_c &= \begin{bmatrix} 18788 & -12100 \\ -2.214 \cdot 10^{11} & -1.427 \cdot 10^{11} \end{bmatrix} \\
B_c &= \begin{bmatrix} -4.146 \\ -3.681 \end{bmatrix} \\
C_c &= \begin{bmatrix} -2.214 \cdot 10^{11} & -1.427 \cdot 10^{11} \end{bmatrix} \\
D_c &= 0.02664 \\
P &= \begin{bmatrix} 1.743 & -0.9675 & 1.743 & -0.9675 \\ -0.9675 & 1.082 & -0.9675 & 1.082 \\ 1.743 & -0.9675 & 10322 & 6648 \\ -0.9675 & 1.082 & 6648 & 4287 \end{bmatrix}
\end{aligned}
$$

for which $\mathrm{spec}(A_c) = \{-6.835, -1.427 \cdot 10^{11}\}$. Due to overly fast dynamics, the resulting controller is in not suitable either for real implementations or numerical simulations. On the other hand, the above controller gives $\mathrm{trace}(X + Y) \approx 10.5569$, *i.e.*, the improvement in terms of suboptimal value is only about 0.017%. Namely, the addition of the above pole placement constraint provides a valuable strategy to design an implementable controller, without penalizing the considered optimization.

Example 3.3. Consider the system derived from [49], already considered in Example 3.1 and that is defined by the following data:

$$
A = \begin{bmatrix} 0 & 1 \\ 0.5 & 0.5 \end{bmatrix}, B = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, K = \begin{bmatrix} -0.3491 & -0.7022 \end{bmatrix}, C = \begin{bmatrix} 0 & 1 \end{bmatrix}, \Delta = 1.
$$

We want to solve (3.60) via Algorithm 3.2. In particular, to start such an algorithm, we use the dynamic output feedback controller issued from the observer-based controller considered in Example 3.1. Namely, by considering the observer gain obtained in the last part of Example 3.1 (the more convenient in terms of the considered optimization), one gets the following data for the initializing controller

$$
\begin{aligned}
A_c &= \begin{bmatrix} -200.7 & 0.2978 \\ -200.2 & -0.2022 \end{bmatrix}, B_c = \begin{bmatrix} 200.4 \\ 200.4 \end{bmatrix}, C_c = \begin{bmatrix} -0.3491 & -0.7022 \end{bmatrix} \\
D_c &= 0.
\end{aligned}
$$

Algorithm 3.2 initialized with the above controller yields

$$
\begin{aligned}
A_c &= \begin{bmatrix} -39.69 & -7.276 \\ -9.156 & -42.67 \end{bmatrix} \\
B_c &= \begin{bmatrix} 0.1338 \\ -0.1285 \end{bmatrix} \\
C_c &= \begin{bmatrix} -28.14 & 82.58 \end{bmatrix} \\
D_c &= -3.734 \\
P &= \begin{bmatrix} 1.575 & -1.159 & 1.575 & -1.159 \\ -1.159 & 1.048 & -1.159 & 1.048 \\ 1.575 & -1.159 & 781.5 & 165.6 \\ -1.159 & 1.048 & 165.6 & 822.2 \end{bmatrix}.
\end{aligned}
$$

Figure 3.3 shows the evolution of the plant state obtained by considering the two different designs. In both simulations, the closed-loop system is initialized as $(x_0, x_c) = (-6, 0, 0, 0)$. Simulations show that the controller designed via the proposed apparatus allows to steer the plant state closer to the origin than for the previously considered observer-based controller.

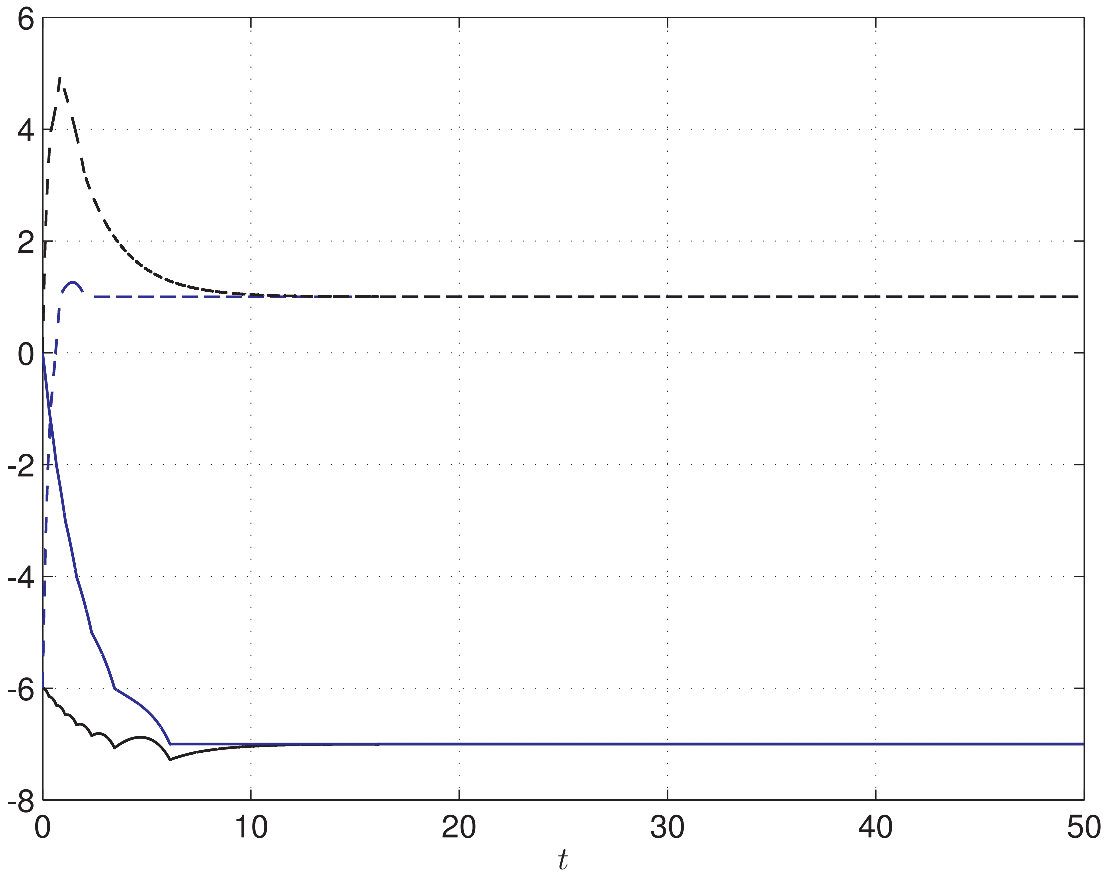


Figure 3.3: Plant state evolution: Proposed design (x_1 dashed-black, x_2 dashed-blue), observer-based control (x_1 solid-black x_2 solid-blue).) The solutions are obtained by integrating the closed-loop model via an Euler method with time step 10^{-4}.

Example 3.4. Consider the linearized model of the Furuta pendulum [67], examined in Example 2.1, for which

$$\dot{x} = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 39.32 & -14.52 & 0 \\ 0 & 81.78 & -13.98 & 0 \end{bmatrix} x + \begin{bmatrix} 0 \\ 0 \\ 25.54 \\ 24.59 \end{bmatrix} u \tag{3.65}$$

where x_1, x_2 represent respectively the base angle and the pendulum angle (rad), x_3 and x_4 are respectively the two angular speeds (rad s^{-1}), and u is input voltage (V) of the motor driving the base shaft. Assume that the two angles x_1, x_2 are measured through two identical incremental optical encoders with resolution of 1°. This situation can be modeled in our setting by taking as measured output $y_m = \mathrm{q}(Cx)$, where

$$C = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}$$

and q is the uniform quantizer defined in (1.17), with $\Delta = \pi/180$. Now, we want to design a dynamic output feedback controller for the given plant by solving (3.60) via Algorithm 3.2. To this end, to initialize Algorithm 3.2, we consider the following LQG controller for the triple (A, B, C)

$$\begin{aligned}
A_c &= \begin{bmatrix} -1.638 & -3.008 & 1 & 0 \\ -3.008 & -14.49 & 0 & 1 \\ 20.17 & -622.7 & 38.48 & -93.35 \\ 1.85 & -639.8 & 37.05 & -89.88 \end{bmatrix} \\
B_c &= \begin{bmatrix} 1.638 & 3.008 \\ 3.008 & 14.49 \\ 5.366 & 25.78 \\ 22.74 & 109 \end{bmatrix} \\
C_c &= \begin{bmatrix} 1 & -24.91 & 2.075 & -3.655 \end{bmatrix} \\
D_c &= \mathbf{0}.
\end{aligned}$$

Algorithm 3.2 initialized with the above controller yields[4]

$$A_c = \begin{bmatrix} -10.22 & -263.1 & 6.119 & -0.03099 \\ 0.1357 & -28.23 & -0.7419 & 3.407 \\ 21.84 & 50.83 & 43.45 & -109.0 \\ 17.94 & -84.79 & 42.89 & -102.5 \end{bmatrix}$$

$$B_c = \begin{bmatrix} -0.6291 & 2.777 \\ -0.0545 & 1.916 \\ -0.6765 & 106.8 \\ -0.3055 & 44.16 \end{bmatrix}$$

$$C_c = \begin{bmatrix} -0.5653 & 9.917 & -1.915 & 2.717 \end{bmatrix}$$

$$D_c = \begin{bmatrix} -0.05517 & -9.548 \end{bmatrix}$$

$$P = \begin{bmatrix} 0.4358 & 0.1078 & 0.4882 & -0.3913 & 0.4358 & 0.1078 & 0.4882 & -0.3913 \\ 0.1078 & 22.54 & 2.812 & -1.624 & 0.1078 & 22.54 & 2.812 & -1.624 \\ 0.4882 & 2.812 & 1.488 & -1.185 & 0.4882 & 2.812 & 1.488 & -1.185 \\ -0.3913 & -1.624 & -1.185 & 1.193 & -0.3913 & -1.624 & -1.185 & 1.193 \\ 0.4358 & 0.1078 & 0.4882 & -0.3913 & 627.1 & -8022 & 644.1 & -1248 \\ 0.1078 & 22.54 & 2.812 & -1.624 & -8022 & 1.044 \cdot 10^5 & -8382 & 16244 \\ 0.4882 & 2.812 & 1.488 & -1.185 & 644.1 & -8382 & 688.5 & -1339.0 \\ -0.3913 & -1.624 & -1.185 & 1.193 & -1248 & 16244 & -1339 & 2610 \end{bmatrix}.$$

Notice that $\lambda_{\max}(P) \approx 108199$, while $\lambda_{\min}(P) \approx 0.0209$. In particular, it is interesting to notice that the eigenvector associated to $\lambda_{\max}(P)$

$$(2.346 \cdot 10^{-7}, -2 \cdot 10^{-4}, -2.243 \cdot 10^{-5}, 1.192 \cdot 10^{-5}, 0.075, -0.98, 0.078, -0.15)$$

is "nearly" parallel to the hyperplane $\mathrm{span}\{e_5, e_6, e_7, e_8\}$, where e_i is the generic vector of the standard basis of $\mathbb{R}^8$, that is the subspace of the state space associated to the controller state. Loosely speaking, the performed optimization, in this case, seems to favor some directions rather than others.

To compare the improvement produced by Algorithm 3.2 with respect to the LQG controller used to initialize such an algorithm, we perform an analysis stage of the two controllers directly employing the conditions provided by Proposition 3.1. Since the measure chosen for the set $\mathcal{A} = \mathcal{E}(P)$ to design the controller is related to $\mathrm{trace}(P^{-1})$, as illustrated in Chapter 2,

[4]A first attempt in the solution to the considered optimization problem leads to a controller unsuitable for physical implementation due to overly fast dynamics. This fact, as already mentioned, can be related to the unattainability of the optima to the considered optimization problem. Thus, in the effective controller design, as already done in the other cases presented in this dissertation, we consider an additional pole placement constraint as the one in (3.63), where α is chosen via a tuning stage aimed at preserving the value of the suboptimal solution obtained.

for each of the two controllers we solve the following optimization problem

$$
\begin{aligned}
\underset{\tau, S_1, S_2, P, \Theta}{\text{minimize}} \quad & \operatorname{trace}(\Theta) \\
\text{subject to} \quad & S_1, S_2 \in \mathcal{D}_+^p, P, \Theta \in \mathcal{S}_+^{2n}, \tau > 0 \\
& \begin{bmatrix} \Theta & \mathrm{I} \\ \bullet & P \end{bmatrix} \geq \mathbf{0}, \ (3.10), (3.11).
\end{aligned}
\tag{3.66}
$$

As usually, to overcome the nonlinearity introduced by the product τP, we perform a grid search for τ. In particular, the solution to the above optimization problem can be performed via an algorithm totally analogous to Algorithm 2.1. By running such an algorithm for the two considered controllers, one gets the following values for $\operatorname{trace}(P)^{-1}$, for the designed controller ($\operatorname{trace}(P_d^{-1})$) and for the LQG controller ($\operatorname{trace}(P_{lqg}^{-1})$)

$$
\begin{aligned}
\operatorname{trace}(P_d^{-1}) &\approx 23.29 \\
\operatorname{trace}(P_{lqg}^{-1}) &\approx 37.6.
\end{aligned}
$$

Namely the proposed design produces an improvement of about 38% with respect to a standard design. Moreover, this improvement in terms of $\operatorname{trace}(P^{-1})$ shows the effectiveness of the implicit minimization of this latter objective via the minimization of $\operatorname{trace}(X + Y)$ performed throughout the design stage.

3.3 Simultaneous Sensor-actuator Quantization

3.3.1 Preliminary Results and Problem Statement

Consider the following continuous-time linear system with sensor and actuator quantization

$$\begin{cases} \dot{x} = Ax + Bv \\ v = \mathrm{q}_u(u) \\ y_m = \mathrm{q}_y(Cx) \end{cases} \tag{3.67}$$

where $x \in \mathbb{R}^n$, $u \in \mathbb{R}^m$, $y_m \in \mathbb{R}^p$, are respectively the state, the input, and the measured output of the plant. A, B, C are real matrices of suitable dimensions, and $\mathrm{q}_u(\cdot), \mathrm{q}_y(\cdot)$ are the uniform quantizers defined in (1.17) having as a quantization error bound, respectively, $\Delta_u, \Delta_y > 0$. We want to design the following plant-order strictly proper dynamic output feedback stabilizing controller for (3.67)

$$\begin{cases} \dot{x}_c = A_c x_c + B_c u_c \\ y_c = C_c x_c \end{cases} \tag{3.68}$$

where $x_c \in \mathbb{R}^n$ is the controller state, $y_c \in \mathbb{R}^m$ is the controller output, $u_c \in \mathbb{R}^p$ is the controller input.

$$(A_c, B_c, C_c) \in \mathbb{R}^{n \times n} \times \mathbb{R}^{n \times p} \times \mathbb{R}^{m \times n}$$

are real matrices to be designed. By interconnecting plant (3.67), *i.e.*, setting $u_c = y_m$, $u = y_c$, with controller (3.68) yields the following dynamics for the closed-loop system

$$\begin{cases} \dot{x} = Ax + B\,\mathrm{q}_u(C_c x_c) \\ \dot{x}_c = A_c x_c + B_c\,\mathrm{q}_y(Cx). \end{cases} \tag{3.69}$$

Remark 3.8. Notice that, the use of a nonstrictly proper controller in this setting induces a nested discontinuity in the closed-loop system; this approach is considered in [37]. However, from a technical point of view, addressing such a nested discontinuity requires a special care. Indeed, in the presence of a nested discontinuity, Proposition 1.2 is of any help. Thus, one needs to extend the results presented in Lemma 2.2 to the case of a composition of the function Ψ with a discontinuous function. On the one hand, such an extension is technically tedious and does not provide any substantial novelty to the proposed methodology. On the other hand, assuming a strictly proper controller does not introduce any severe restriction in the apparatus presented in the sequel. Therefore, to maintain the presentation simple and to focus more on the key ideas and results, we will insist in the remainder of this chapter in considering a strictly proper controller.

Therefore, by defining the functions

$$\begin{aligned}\Psi_u\colon \mathbb{R}^m &\to \mathbb{R}^m\\ z &\mapsto \mathrm{q}_u(z) - z\end{aligned} \tag{3.70a}$$

$$\begin{aligned}\Psi_y\colon \mathbb{R}^p &\to \mathbb{R}^p\\ z &\mapsto \mathrm{q}_y(z) - z\end{aligned} \tag{3.70b}$$

by taking as vector state $\tilde{x} = (x, x_c) \in \mathbb{R}^{2n}$, and by defining the matrices

$$\widetilde{A} = \begin{bmatrix} A & BC_c \\ B_cC & A_c \end{bmatrix}, \widetilde{B}_1 = \begin{bmatrix} \mathbf{0} \\ B_c \end{bmatrix}, \widetilde{B}_2 = \begin{bmatrix} B \\ \mathbf{0} \end{bmatrix}, \widetilde{C}_1 = \begin{bmatrix} C & \mathbf{0} \end{bmatrix}, \widetilde{C}_2 = \begin{bmatrix} \mathbf{0} & C_c \end{bmatrix} \tag{3.71}$$

(3.69) can be rewritten as

$$\dot{\tilde{x}} = \widetilde{A}\tilde{x} + \widetilde{B}_1\Psi_y\left(\widetilde{C}_1\tilde{x}\right) + \widetilde{B}_2\Psi_u\left(\widetilde{C}_2\tilde{x}\right). \tag{3.72}$$

At this stage, since the functions Ψ_u, Ψ_y are discontinuous, the right-hand side of (3.72) is a discontinuous function of the state. Thus, we want to focus on Krasovskii solutions to system (3.72). In view of the local boundedness of the right-hand side of (3.72), for every $\tilde{x}_0 \in \mathbb{R}^{2n}$, there exists at least a Krasovskii solution φ to (3.72) with $\varphi(0) = \tilde{x}_0$; (see Chapter 1). In particular, let us define

$$\begin{aligned}X\colon \mathbb{R}^{2n} &\to \mathbb{R}^{2n}\\ \tilde{x} &\mapsto \widetilde{A}\tilde{x} + \widetilde{B}_1\Psi_y(\widetilde{C}_1\tilde{x}) + \widetilde{B}_2\Psi_u(\widetilde{C}_2\tilde{x})\end{aligned} \tag{3.73a}$$

we consider the solutions to the following differential inclusion

$$\dot{\tilde{x}} \in \mathcal{K}[X](\tilde{x}) \tag{3.73b}$$

where $\mathcal{K}[X](\tilde{x})$ represents the Krasovskii regularization of the function X; see Definition 1.2 on Page 14. The next theorem provides a first characterization of the behavior of (3.73).

Theorem 3.2. Let A, B, C, A_c, B_c, C_c be matrices of adequate dimensions such that $\widetilde{A}$ is Hurwitz. Then, there exists a compact set $\mathcal{A} \subset \mathbb{R}^{2n}$, containing the origin, which is UGAS for (3.73).

Proof. The proof of the above result follows the same lines of the proof of Theorem 2.1. In particular, under the considered hypothesis, we derive for (3.73) a relation like (2.8). Then, the proof directly follows from the arguments presented in the proof of Theorem 2.1.

For every $\tilde{x} \in \mathbb{R}^{2n}$, define $c_1(\tilde{x}) = \widetilde{C}_1\tilde{x}$ and $c_2(\tilde{x}) = \widetilde{C}_2\tilde{x}$. Since the function $\tilde{x} \mapsto \widetilde{A}\tilde{x}$ is continuous, by Proposition 1.1, for every $\tilde{x} \in \mathbb{R}^{2n}$,

$$\mathcal{K}[X](\tilde{x}) = \widetilde{A}\tilde{x} + \mathcal{K}[\widetilde{B}_1\Psi_y \circ c_1 + \widetilde{B}_2\Psi_u \circ c_2](\tilde{x})$$

Therefore, by item ii of Proposition 1.1, for each $\tilde{x} \in \mathbb{R}^{2n}$, it follows that

$$\mathcal{K}[\widetilde{B}_1\Psi_y \circ c_1 + \widetilde{B}_2\Psi_u \circ c_2](\tilde{x}) \subseteq \widetilde{B}_1\mathcal{K}[\Psi_y \circ c_1](x) + \widetilde{B}_2\mathcal{K}[\Psi_u \circ c_2](x). \tag{3.74}$$

Moreover, due to the bound shown earlier on the function Ψ, it turns out that for each $\tilde{x} \in \mathbb{R}^{2n}$

$$\mathcal{K}[\Psi_y \circ c_1](\tilde{x}) \subseteq \mathbb{B}\sqrt{p}\Delta_y$$
$$\mathcal{K}[\Psi_u \circ c_2](\tilde{x}) \subseteq \mathbb{B}\sqrt{m}\Delta_u$$

therefore, for every $\tilde{x} \in \mathbb{R}^{2n}$, the following inclusion holds

$$\mathcal{K}[X](\tilde{x}) \subseteq \widetilde{A}\tilde{x} + \widetilde{B}_1\mathbb{B}\sqrt{p}\Delta_y + \widetilde{B}_2\mathbb{B}\sqrt{m}\Delta_u. \tag{3.75}$$

Since $\widetilde{A}$ is Hurwitz there exist $P, Q \in \mathcal{S}_+^{2n}$, such that $\mathrm{He}(P\widetilde{A}) = -Q$. Building on this relation, for each $\tilde{x} \in \mathbb{R}^{2n}$, define the function $V(\tilde{x}) = \tilde{x}^\mathsf{T} P\tilde{x}$. Then, thanks to (3.75), for every $\tilde{x} \in \mathbb{R}^{2n}$, and any $f \in \mathcal{K}[X](\tilde{x})$

$$\langle \nabla V(x), f\rangle = -\tilde{x}^\mathsf{T} Q\tilde{x} + 2\tilde{x}^\mathsf{T} P\left(\widetilde{B}_1\xi_y + \widetilde{B}_2\xi_u\right)$$

for some $\xi_y \in \mathbb{B}\sqrt{p}\Delta_y, \xi_u \in \mathbb{B}\sqrt{m}\Delta_u$. By standard arguments, it is straightforward to show that

$$\|P\widetilde{B}_1\xi_y + P\widetilde{B}_2\xi_u\|^2 \leq 2\left(\|P\widetilde{B}_1\xi_y\|^2 + \|P\widetilde{B}_2\xi_u\|^2\right) \leq 2\left(\|P\widetilde{B}_1\|^2\Delta_y^2 p + \|P\widetilde{B}_2\|^2\Delta_u^2 m\right).$$

Thus, by following the same arguments shown in the proof of Theorem 2.1, pick $\theta \in (0, 1)$, then

$$\mathcal{A} = \left\{\tilde{x} \in \mathbb{R}^{2n} \colon V(\tilde{x}) \leq \frac{8\lambda_{\max}(P)}{\lambda_{\min}^2(Q)\theta}\left(\|P\widetilde{B}_1\|^2\Delta_y^2 p + \|P\widetilde{B}_2\|^2\Delta_u^2 m\right)\right\}$$

is UGAS for (3.73), concluding the proof. ■

Building on the above result, as already done throughout this dissertation, with the aim of providing constructive conditions for the design of the controller (3.68) ensuring UGAS of a certain compact set, we want to derive sufficient conditions solving the problem given next.

Problem 3.4. Let A, B, C be matrices of adequate dimensions. Determine $(A_c, B_c, C_c) \in \mathbb{R}^{n\times n} \times \mathbb{R}^{n\times p} \times \mathbb{R}^{m\times n}$, and a compact set $\mathcal{A} \subset \mathbb{R}^{2n}$ containing the origin, such that $\mathcal{A}$ is UGAS for system (3.73).

3.3.2 Sufficient Conditions

A first sufficient condition to solve Problem 3.4, and based on the sector conditions illustrated in Lemma 2.2, is given next.

Proposition 3.5. *If there exist $P \in \mathcal{S}_+^{2n}$, $S_1, S_2 \in \mathcal{D}_+^p$, $\bar{S}_1, \bar{S}_2 \in \mathcal{D}_+^m$, A_c, B_c, C_c real matrices* *of adequate dimensions, and a positive scalar τ such that*

$$\begin{bmatrix} \mathrm{He}(P\widetilde{A})+\tau P & P\widetilde{B}_1-\widetilde{C}_1^{\mathsf{T}}S_2 & P\widetilde{B}_2-\widetilde{C}_2^{\mathsf{T}}\bar{S}_2 \\ \bullet & -2S_2-S_1 & \mathbf{0} \\ \bullet & \bullet & -2\bar{S}_2-\bar{S}_1 \end{bmatrix} < \mathbf{0} \tag{3.76}$$

$$\mathrm{trace}(S_1)\Delta_y^2+\mathrm{trace}(\bar{S}_1)\Delta_u^2-\tau\leq 0 \tag{3.77}$$

where

$$\widetilde{A}=\begin{bmatrix} A & BC_c \\ B_cC & A_c \end{bmatrix}, \widetilde{B}_1=\begin{bmatrix} \mathbf{0} \\ B_c \end{bmatrix}, \widetilde{B}_2=\begin{bmatrix} B \\ \mathbf{0} \end{bmatrix}, \widetilde{C}_1=\begin{bmatrix} C & \mathbf{0} \end{bmatrix}, \widetilde{C}_2=\begin{bmatrix} \mathbf{0} & C_c \end{bmatrix}.$$

then

$$A_c, B_c, C_c \tag{3.78}$$

$$\mathcal{A}=\mathcal{E}(P) \tag{3.79}$$

solve Problem 3.1.

Proof. For every $\tilde{x}\in\mathbb{R}^{2n}$, consider the following quadratic function $V(\tilde{x})=\tilde{x}^{\mathsf{T}}P\tilde{x}$, where $P\in\mathcal{S}_+^{2n}$. Following the ideas presented in the proof of Theorem 2.1, we want to prove that there exists a positive real scalar β such that

$$\langle\nabla V(\tilde{x}),w\rangle\leq-\beta V(\tilde{x}) \qquad \forall\tilde{x}\in\mathbb{R}^{2n}\setminus\mathrm{Int}\mathcal{A}, w\in\mathcal{K}[X](\tilde{x}). \tag{3.80}$$

As the above relation is analogous to (2.8) in the proof of Theorem 2.1, establishing (3.80) suffices to show that the set $\mathcal{A}$ in (3.79) is UGAS for (3.73). By S-procedure arguments, (3.80) can be verified by showing that for every $\tilde{x}\in\mathbb{R}^{2n}$, there exists a positive real scalar τ such that

$$\langle\nabla V(\tilde{x}),w\rangle-\tau(1-\tilde{x}^{\mathsf{T}}P\tilde{x})\leq-\beta V(\tilde{x}) \qquad \forall w\in\mathcal{K}[X](\tilde{x}). \tag{3.81}$$

On the other hand, as shown in the proof of Theorem 3.2, for every $w\in\mathcal{K}[X](\tilde{x})$, there exist $v_1\in\mathcal{K}[\Psi_y](\widetilde{C}_1\tilde{x})$ and $v_2\in\mathcal{K}[\Psi_u](\widetilde{C}_2\tilde{x})$, such that $w=\widetilde{A}\tilde{x}+\widetilde{B}_1v_1+\widetilde{B}_1v_2$. Then, still by S-procedure arguments and according to Lemma 2.2, (3.81) is ensured by proving that for each $\tilde{x}\in\mathbb{R}^{2n}$, and for each $v_1\in\mathbb{R}^p$,$v_2\in\mathbb{R}^m$

$$\begin{aligned}&\langle\nabla V(\tilde{x}),\widetilde{A}\tilde{x}+\widetilde{B}_1v_1+\widetilde{B}_2v_2\rangle-\tau(1-\tilde{x}^{\mathsf{T}}P\tilde{x})-v_1^{\mathsf{T}}S_1v_1+\mathrm{trace}(S_1)\Delta_y^2-2v_1^{\mathsf{T}}S_2(v_1+\widetilde{C}_1\tilde{x})\\&-v_2^{\mathsf{T}}\bar{S}_1v_2+\mathrm{trace}(\bar{S}_1)\Delta_u^2-2v_2^{\mathsf{T}}\bar{S}_2(v_2+\widetilde{C}_2\tilde{x})\leq-\beta V(\tilde{x}).\end{aligned} \tag{3.82}$$

By straightforward calculations, the left-hand side of the above relation can be rewritten as follows

$$\begin{bmatrix}\tilde{x}\\ v_1\\ v_2\end{bmatrix}^T \begin{bmatrix}\mathrm{He}(P\widetilde{A})+\tau P & P\widetilde{B}_1-\widetilde{C}_1^{\mathsf{T}}S_2 & P\widetilde{B}_2-\widetilde{C}_2^{\mathsf{T}}\bar{S}_2\\ \bullet & -2S_2-S_1 & \mathbf{0}\\ \bullet & \bullet & -2\bar{S}_2-\bar{S}_1\end{bmatrix}\begin{bmatrix}\tilde{x}\\ v_1\\ v_2\end{bmatrix} + \mathrm{trace}(S_1)\Delta_y^2+\mathrm{trace}(\bar{S}_1)\Delta_u^2-\tau. \tag{3.83}$$

Thus in view of (3.76) and (3.77), it follows that there exists a small enough positive scalar γ such that for every $\tilde{x}\in\mathbb{R}^{2n}\setminus\mathrm{Int}\mathcal{A}, w\in\mathcal{K}[X](\tilde{x})$, one has $\langle\nabla V(x),w\rangle\leq-\gamma\tilde{x}^{\mathsf{T}}\tilde{x}$. Then, since for every $\tilde{x}\in\mathbb{R}^{2n}$, $V(\tilde{x})\leq\lambda_{\max}(P)\tilde{x}^{\mathsf{T}}\tilde{x}$, by setting $\beta=\frac{\gamma}{\lambda_{\max}(P)}$ gives (3.81), and this finishes the proof. ■

Also in this case, the above result is lossless with respect to Theorem 3.2. Precisely, Assumption 1.2 ensures the feasibility of conditions (3.76) and (3.77). This claim is formalized in the result given next.

Proposition 3.6. *Let A,B,C matrices such that Assumption 1.2 is satisfied. Then, there exist*

$$(\tau,P,S_1,S_2,\bar{S}_1,\bar{S}_2,A_c,B_c,C_c)\in\mathbb{R}_{>0}\times\mathcal{S}_+^{2n}\times\mathcal{D}_+^p\times\mathcal{D}_+^p\times\mathcal{D}_+^m\times\mathcal{D}_+^m\times\mathbb{R}^{n\times n}\times\mathbb{R}^{n\times p}\times\mathbb{R}^{m\times n}$$

satisfying (3.76) and (3.77).

Proof. Define

$$\widehat{B}=\begin{bmatrix}\widetilde{B}_1 & \widetilde{B}_2\end{bmatrix}, \widehat{C}=\begin{bmatrix}\widetilde{C}_1\\ \widetilde{C}_2\end{bmatrix}, \widehat{S}_1=\mathrm{diag}\{S_1,\bar{S}_1\}, \widehat{S}_2=\mathrm{diag}\{S_2,\bar{S}_2\}$$

then, (3.76) can be equivalently rewritten as

$$\begin{bmatrix}\mathrm{He}(P\widetilde{A})+\tau P & P\widehat{B}-\widehat{C}^{\mathsf{T}}\widehat{S}_2\\ \bullet & -2\widehat{S}_2-\widehat{S}_1\end{bmatrix}<\mathbf{0}. \tag{3.84}$$

Now, observe that for each $S_1\in\mathcal{D}_+^p, \bar{S}_1\in\mathcal{D}_+^m$, one has

$$\mathrm{trace}(S_1)\Delta_y^2+\mathrm{trace}(\bar{S}_1)\Delta_u^2\leq\mathrm{trace}(\widehat{S}_1)\max(\Delta_y^2,\Delta_u^2).$$

Therefore, if there exists

$$(\tau,P,S_1,S_2,\bar{S}_1,\bar{S}_2,A_c,B_c,C_c)\in\mathbb{R}_{>0}\times\mathcal{S}_+^{2n}\times\mathcal{D}_+^p\times\mathcal{D}_+^p\times\mathcal{D}_+^m\times\mathcal{D}_+^m\times\mathbb{R}^{n\times n}\times\mathbb{R}^{n\times p}\times\mathbb{R}^{m\times n}$$

such that (3.84) and

$$\mathrm{trace}(\widehat{S}_1)\max(\Delta_y^2,\Delta_u^2)-\tau\leq 0 \tag{3.85}$$

are satisfied, so are (3.76) and (3.77). On the other hand, since Assumption 1.2 ensures the existence of A_c,B_c,C_c such that $\widetilde{A}$ is Hurwitz, and (3.84) and (3.85), respectively, match (2.14) and (2.15), by following the same arguments proposed in the proof of Proposition 2.2, the assert is proven. ■

3.3.3 Controller Design

Also in this case, the conditions issued from Proposition 3.1 cannot be directly employed to solve Problem 3.4. On the one hand, attempting to design an observer-based control gives rise to the same drawbacks discussed in the previous section with even an increased complexity due to the addition of actuator quantization. On the other hand, the similarities between the conditions in Proposition 3.5 with the ones in Proposition 3.1 foster to reconsider the same approach pursued in Proposition 3.4. In particular, retracing the steps performed to derive Proposition 3.4 gives rise to the following result.

Proposition 3.7. *For each τ, $S_1, S_2 \in \mathcal{D}_+^p$, $\bar{S}_1, \bar{S}_2 \in \mathcal{D}_+^m$ there exist $K \in \mathbb{R}^{n\times n}$, $L \in \mathbb{R}^{n\times p}$, $X, Y \in \mathcal{S}_+^n$, and $M \in \mathbb{R}^{m\times n}$ such that*

$$\begin{bmatrix} \mathrm{He}(H_1) + \tau H_2 & H_3 & H_4 \\ \bullet & -S_1 - 2S_2 & \mathbf{0} \\ \bullet & \bullet & -\bar{S}_1 - 2\bar{S}_2 \end{bmatrix} < \mathbf{0} \tag{3.86}$$

$$H_2 > \mathbf{0} \tag{3.87}$$

where

$$H_1 = \begin{bmatrix} AY + BM & A \\ K & XA + LC \end{bmatrix}, H_2 = \begin{bmatrix} Y & \mathrm{I} \\ \bullet & X \end{bmatrix}, H_3 = \begin{bmatrix} -YC^{\mathsf{T}} S_2 \\ L - C^{\mathsf{T}} S_2 \end{bmatrix}, H_4 = \begin{bmatrix} B - M^{\mathsf{T}} \bar{S}_2 \\ XB - \bar{S}_2 \end{bmatrix}$$

if and only if, for any nonsingular matrices $U, V \in \mathbb{R}^{n\times n}$ such that $UV^{\mathsf{T}} = \mathrm{I} - XY$, $\hat{X} = U^{\mathsf{T}}(X - Y^{-1})^{-1}U$,

$$\begin{aligned} C_c &= MV^{-\mathsf{T}} \\ B_c &= U^{-1}L \\ A_c &= U^{-1}\left(K - XAY - XBM - UB_cCY\right)V^{-\mathsf{T}} \end{aligned} \tag{3.88}$$

and

$$P = \begin{bmatrix} X & U \\ \bullet & \hat{X} \end{bmatrix} \tag{3.89}$$

satisfy (3.76).

Proof. The proof is totally analogous to the proof of Proposition 3.4. In particular, necessity can be proven as in the proof of Proposition 3.4 by still employing the same change of variables as in (3.59), with the only caveat to enforce $D_c = \mathbf{0}$, and by noticing that

$$\mathbb{J}PB_2 = \begin{bmatrix} B \\ XB \end{bmatrix}, \mathbb{J}\widetilde{C}_2^{\mathsf{T}} = \begin{bmatrix} -M^{\mathsf{T}} \\ \mathrm{I} \end{bmatrix}.$$

Sufficiency can be proven directly by retracing the same steps as in the proof of Proposition 3.4, with the only caveat to enforce $N = \mathbf{0}$. ∎

3.3.4 Optimization and Numerical Issues

Also for the matter of Problem 3.4, we want to associate to such a problem the following optimization problem

Problem 3.5 (Stabilization)**.** Let A, B, C be matrices of adequate dimensions.
Determine A_c, B_c, C_c, and $P \in \mathcal{S}_+^{2n}$, such that $\mathcal{E}(P)$ is UGAS for system (3.73) and it is minimized with respect to some criterion.

Pursuing the same approach shown in the first section of this chapter, which relies on Proposition 3.7, Problem 3.5 turns into

$$\begin{aligned} &\underset{X,Y,L,K,M,\tau,S_1,S_2,\bar{S}_1,\bar{S}_2}{\text{minimize}} && \operatorname{trace}(X+Y) \\ &\text{subject to} && S_1, S_2 \in \mathcal{D}_+^p, \bar{S}_1, \bar{S}_2 \in \mathcal{D}_+^m, X, Y \in \mathcal{S}_+^n, \tau > 0 \\ & && (3.86), (3.87), (3.77). \end{aligned} \tag{3.90}$$

Obviously the solution to (3.90) entails the same issues discussed on Page 102, with an increased complexity due to the further nonlinearity introduced by the bilinear term $M^\mathsf{T}\bar{S}_2$, and its symmetric, appearing in (3.86) due to actuator quantization. Nevertheless, the same strategies presented in the previous section can be adopted to face this problem. On the one hand, a grid search in this setting entails a greater number of elements subject to such a search, namely $p+m+1$. On the other hand, the complexity of an iterative procedure as the one presented in Algorithm 3.2 is unchanged since $S_2, \bar{S}_2$ are selected once at the same time throughout the first step. Moreover, in light of Proposition 3.6 feasibility of the optimization problems considered at each step can be ensured via suitable choice. Thus, the adoption of an algorithm alike to Algorithm 3.2 is certainly a viable solution to tackle (3.90).

Numerical Example

Example 3.5. Let us consider again the system analyzed in Example 3.4, and assume that, in addition to sensor quantization as considered in Example 3.4, the plant is subject to uniform actuator quantization with quantization error bound $\Delta_u = 0.25$. That situation can be embedded in the setting illustrated in (3.67), by taking $\Delta_y = \pi/180$, $\Delta_u = 0.25$. To stabilize the closed-loop system, we want to design the controller in (3.68) by solving (3.90) via an algorithm totally analogous to Algorithm 3.2. The initialization of such an algorithm is performed choosing as initializing controller the LQG controller already considered in

Example 3.4. In particular, under this choice, the mentioned algorithm yields

$$
\begin{aligned}
A_c &= \begin{bmatrix} -70.5 & 241.5 & -44.31 & 57.22 \\ -4.029 & -34.87 & -0.169 & 2.32 \\ -18.91 & -1536 & 40.06 & -148 \\ 22.47 & -1736 & 69.26 & -178.9 \end{bmatrix} \\
B_c &= \begin{bmatrix} 22.22 & -170.4 \\ -4.091 & -40.34 \\ -11.16 & -548 \\ -2.971 & -557 \end{bmatrix} \\
C_c &= \begin{bmatrix} -5.065 & 84.51 & -7.082 & 11.84 \end{bmatrix} \\
P &= \begin{bmatrix} 3.785 & -6.101 & -2.43 & 1.689 & 3.785 & -6.101 & -2.43 & 1.689 \\ -6.101 & 59.4 & 6.855 & -6.52 & -6.101 & 59.4 & 6.855 & -6.52 \\ -2.43 & 6.855 & 2.6 & -2.394 & -2.43 & 6.855 & 2.6 & -2.394 \\ 1.689 & -6.52 & -2.394 & 2.465 & 1.689 & -6.52 & -2.394 & 2.465 \\ 3.785 & -6.101 & -2.43 & 1.689 & 6.074 & -20.48 & -0.4997 & -0.133 \\ -6.101 & 59.4 & 6.855 & -6.52 & -20.48 & 163.5 & -6.034 & 6.759 \\ -2.43 & 6.855 & 2.6 & -2.394 & -0.4997 & -6.034 & 4.296 & -4.048 \\ 1.689 & -6.52 & -2.394 & 2.465 & -0.133 & 6.759 & -4.048 & 4.206 \end{bmatrix}
\end{aligned}
$$

As[5] in Example 3.4, to compare the improvement arisen by the use of Algorithm 3.2 with respect to the LQG controller used to initialize such an algorithm, we perform an analysis stage of the two controllers directly employing the conditions provided by Proposition 3.5. Since the measure chosen for the set $\mathcal{A} = \mathcal{E}(P)$ to design the controller is related to trace(P^{-1}), as illustrated in Chapter 2, for each of the two controllers we solve the following optimization problem

$$
\begin{array}{ll}
\underset{\tau, S_1, S_2, P, \Theta}{\text{minimize}} & \text{trace}(\Theta) \\
\text{subject to} & S_1, S_2 \in \mathcal{D}_+^p, P, \Theta, \in \mathcal{S}_+^{2n}, \tau > 0 \\
& \begin{bmatrix} \Theta & \mathrm{I} \\ \bullet & P \end{bmatrix} \geq \mathbf{0}, (3.76), (3.77).
\end{array}
\tag{3.91}
$$

As usually, to overcome the nonlinearity introduced by the product τP, we perform a grid search for τ. In particular, the solution to the above optimization problem can be performed via an algorithm similar to Algorithm 2.1. By running such an algorithm for the two considered controllers, one gets the following values for trace$(P)^{-1}$, for the designed controller

[5] Also in this case, a first attempt in the solution to the considered optimization problem leads to a controller unsuitable for physical implementation due to overly fast dynamics and poorly damped eigenvalues. Thus, in the effective controller design, as already done in the other cases presented in this dissertation, we consider an additional pole placement constraint as those in (3.63) and (3.64) characterized by parameters α and θ chosen via a tuning stage aimed at preserving the value of the suboptimal solution obtained.

(trace(P_d^{-1})) and for the LQG controller (trace(P_{lqg}^{-1}))

$$\mathrm{trace}(P_d^{-1}) \approx 57.01$$
$$\mathrm{trace}(P_{lqg}^{-1}) \approx 855.32$$

That is, the proposed design produces an improvement of about 93.33% with respect to the considered standard LQG design used to initialize the proposed algorithm. Figure 3.4 and Figure 3.5 show, respectively, the steady-state evolution of the plant state and of the controller state obtained by considering the two different controllers. In both simulations, the closed-loop system is initialized as $(x_0, x_c) = (0, \pi/4, 0, 0, \mathbf{0}_4)$. Simulations bring out that the proposed design allows to notably reduce the amplitude of the oscillations induced by quantization.

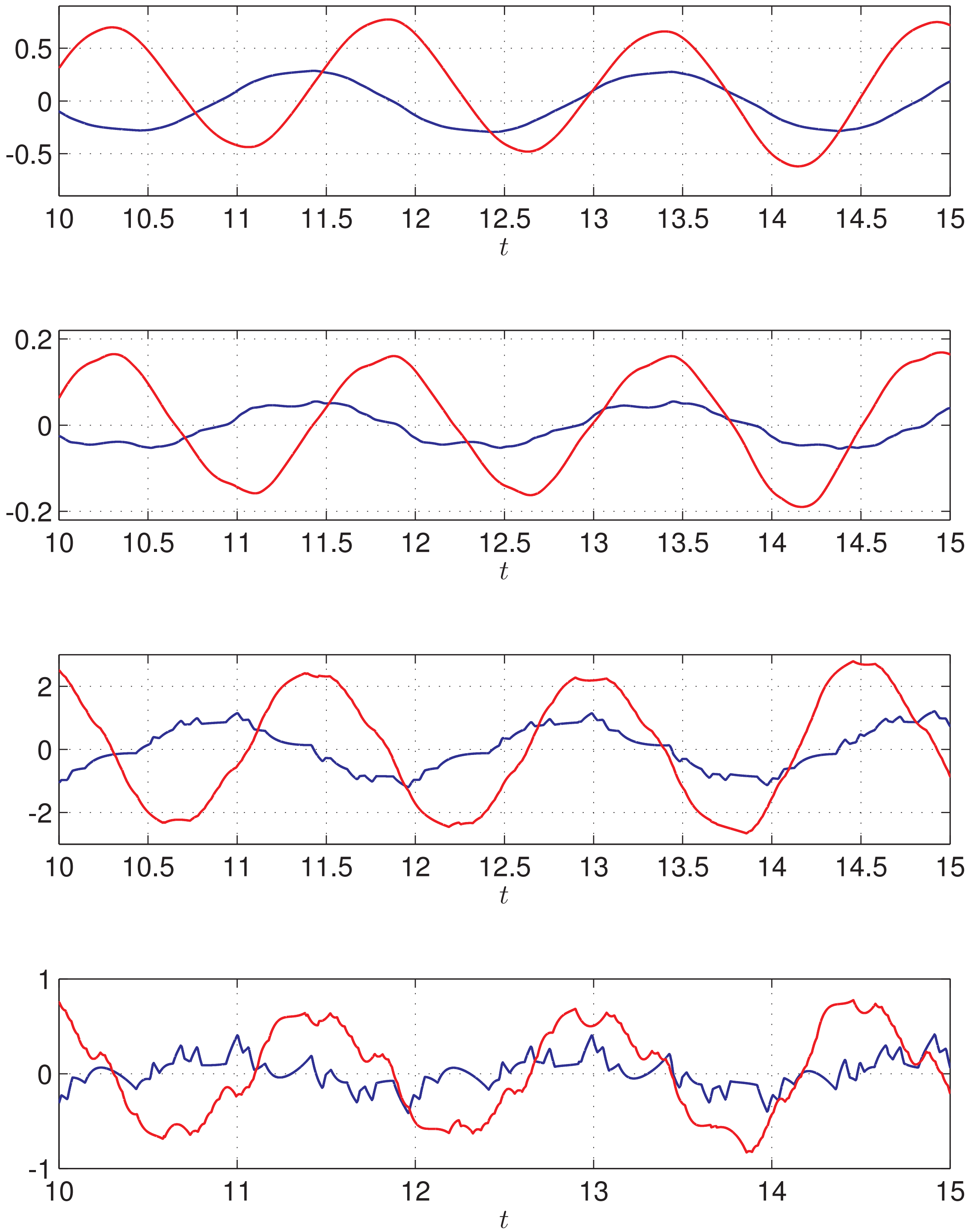


Figure 3.4: Plant state evolution: Proposed design (blue), LQG design (red). The solutions are obtained by integrating the closed-loop model via an Euler method with time step 10^{-4}.

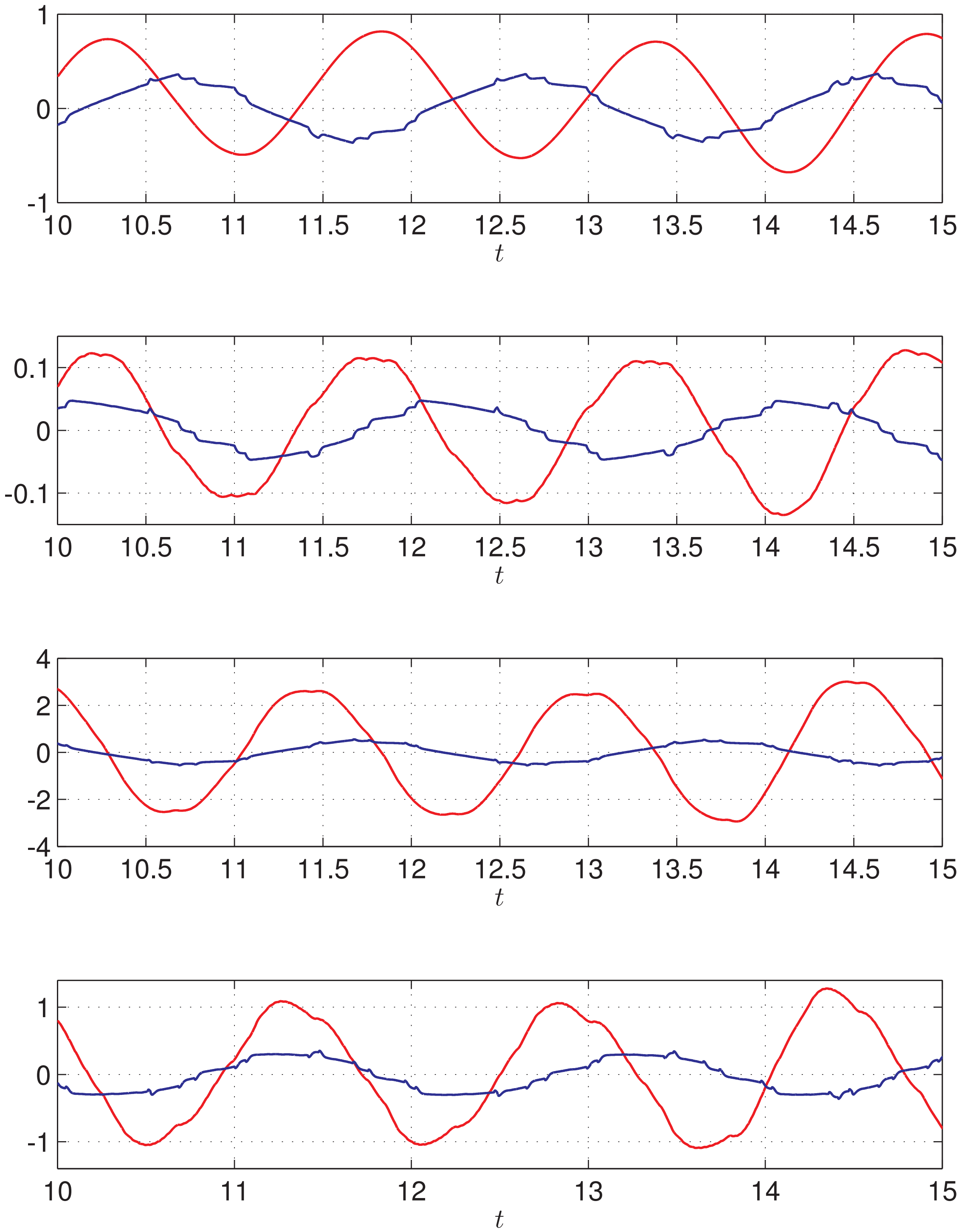


Figure 3.5: Controller state evolution: Proposed design (blue), LQG design (red). The solutions are obtained by integrating the closed-loop model via an Euler method with time step 10^{-4}.

3.4 Comments and Conclusion

In this chapter, we tackled the design problem of a dynamic output feedback controller to stabilize linear plants in the presence of sensor quantization and simultaneous sensor-actuator quantization. In this setting, we firstly shown that Assumption 1.2 ensures the existence of a compact set $\mathcal{A}$ containing the origin, which is UGAS for the quantized control systems, with respect to Krasovskii solutions. Such a result also points out that the compact set $\mathcal{A}$ can be chosen as a sublevel set of a certain quadratic function. Thus, building on this result, by the use of quadratic Lyapunov-like functions coupled via S-procedure to the sector conditions for the uniform quantizer illustrated in Chapter 2, we turned the stabilization problem into the feasibility problem to certain matrix inequalities. Such a formulation based on matrix inequalities is shown to be lossless in the sense that under Assumption 1.2, the derived matrix inequalities are always feasible. Thus, the proposed formulation not only structures the design problem, but also decreases the conservatism in the determination of the set $\mathcal{A}$ with respect to the main result without requiring any additional hypothesis beyond Assumption 1.2.

Afterward, relying on the proposed characterization of the stabilization problem based on matrix inequalities, we proposed a complete apparatus based on convex optimization over LMIs to allow the controller design while shrinking the size of the set $\mathcal{A}$. The effectiveness of the proposed methodology is shown in some examples. As mentioned in the previous chapter of this dissertation, the methodology proposed is quite flexible to envision the extension of the derived results to finite range quantizers, as well as to other kind of quantizers. In particular, for the extension to finite range quantizers, the same considerations discussed in the end of Chapter 2 about finite range quantizers, and other kinds of quantizers apply *mutatis mutandis* for the matter of the problem considered in this chapter.

The results presented in this chapter show that employing an observer-based controller in the presence of sensor quantization does not allow to derive computationally tractable conditions for the design the complete design of the resulting controller. In particular, as shown, one needs first to make a choice for the controller gain K, and then designing the observer gain L via the solution to convex optimization problem over LMIs. Nonetheless, we shown that if the choice considered for K is such that the matrix $A + BK$ is Hurwitz, then the resulting optimization problem allowing the design of the gain L is always feasible. Such a shortcoming preventing from fully designing an output feedback controller resting on an observer-based architecture is completely overcome by considering a general plant-order dynamic output feedback controller. The adoption of the latter controller scheme also allows, with few extra work, to derive computationally tractable conditions for the design of an output feedback controller to deal with simultaneous sensor and actuator quantization, bridging the gap left in Chapter 2. However, it is worthwhile to notice that adopting a dynamic controller entails an augmentation of the closed-loop system state, whose turns out to be the aggregation of the controller state and of the plant state. This fact could lead to unsatisfactory results in terms of the behavior of the plant. Indeed, whenever the stabilization of the closed-loop system pertains to the whole control system state, the considered optimization aimed at steering the closed-loop system state as much as possible to the origin, could favor more the component of the state associated to the controller than the ones associated to the plant, leading to large deviations of the plant state from the origin. For this reason, the paradigm of steering the plant state as much as possible close to the origin via the shrinkage of the set $\mathcal{A}$ naturally considered in Chapter 2 needs to be partially reconsidered in the presence of additional dynamics in the closed-loop system. As shown, this point can be (partially) addressed by considering an observer-based controller architecture. In fact as pointed in Remark 3.1, the adoption of this architecture enables to somehow decouple the optimization to focus more on the side of the plant. However, as underlined, this kind of architecture is hard to manage from a numerical standpoint. A possible solution to overcome this problem consists of considering a size criterion based on directions of interest as the ones considered in Chapter 2 also for the design of the full dynamic controller considered in Section 3.2.4, and Section 3.3.2. Such directions can be chosen to belong to the subspace of the state space associated to the plant state. On the other hand, pursuing this approach would not suggest any selection for the matrix U in Proposition 3.4 and Proposition 3.7. Such a further variable could be considered to shape the issued controller in a way that ensures its physical construction and/or additional requirements. This aspect provides an interesting direction for future research.

This page is intentionally left blank.

CONCLUSION OF PART I

Concluding Remarks

In this part, we provided several tools to perform stability analysis and controller design of linear systems in the presence of actuator and/or sensor quantization. In particular, both static state feedback control and dynamic output feedback controllers were considered. The pursued approach strives to always lead to computationally tractable conditions, so as to provide solid and reliable tools actually exploitable in real-world settings. This latter feature stems from having founded the whole methodology on convex optimization, in particular we proposed an LMI-based approach.

Another interesting feature of our approach consists of having adopted the notion of solutions due to Krasovskii. This choice allows to both overcome the technical issues concerning the existence of solutions for the closed-loop system and to exploit a large class of existing results presented in the literature. In particular, the exploitation of such results allows to certify stronger properties for the solutions to the considered closed-loop systems than the ones usually considered. We emphasize that the analysis we considered takes into account Carathéodory solutions whenever they exist.

Moreover, we would like to point out that having dealt with Krasovskii solutions, due to the equivalence between Krasovskii solutions and Hermes solutions mentioned in Chapter 1, guarantees that the properties established for the closed-loop system are robust with respect to small perturbations, that inevitably affect physical control systems.

Another interesting aspect pertains to the fact that having considered Krasovskii solutions does not lead to any change in the resulting constructive procedures with respect to classical approach. Notice that this aspect is only due to the fact that the sector conditions we worked out for the quantizer considered in this dissertation provide sufficiently room to include the set-valued mapping resulting from the Krasovskii regularization of the closed-loop system.

Observe that, in general, this may not be the case.

We would like to point out that some preliminary results combining quantization, time-delays and saturation have been presented in [39].

As pointed out throughout this first part, the main drawbacks encountered essentially concern the fact that most of the time the approach we followed does not lead straight to genuine convex optimization problems. This shortcoming has been addressed by the introduction of specific iterative algorithms able to handle the optimization problems issuing from the considered problems. The most important features offered by the algorithms we proposed consists of:

- avoiding as much a possible the use of tuning stages and/or heuristics,
- always providing a suboptimal solution to the considered optimization problems.

The two above properties are of primary interest to envision solid and systematic tools to be exploited in real-world applications. On the other hand, such algorithms operate iteratively, hence they may lead to an increased complexity from a numerical standpoint. Moreover, the convergence toward the optimum (whenever it exists) cannot be guaranteed.

Perspectives and Future Outlook

The methodology we offered appears quite robust and promising to envision several extensions. Such extensions, as briefly discussed all along this dissertation up to now, mainly consist of considering finite range quantizers and dealing with other class of quantizers, as the well established finite precision logarithmic quantizer; see [21]. Another possible line of research pertains to the extension of the methodology we proposed to a wider class of plants as for instance polynomial systems. This class of systems has been recently achieving a resounding interest by researchers due to the emerging of solid numerical tools to address a large number of problems originating in such a scenario; see [24, 61]. In this context, an interesting issue lies in generalizing the methodology illustrated in this dissertation via the use of polynomial Lyapunov-like functions instead of quadratic ones.

Although the discontinuous behaviors induced by quantizers are fully accounted by the proposed analysis, such a discontinuity may induce behaviors that are undesired in real control systems. Such behaviors essentially consist of rapid switching experienced by quantized variables. Such a phenomenon is induced by unattainable sliding-mode and/or by the presence of process and measurement disturbances, always present in engineered control systems. On the one hand, these phenomena induce an early wear of physical elements used to the real implementation of control systems. On the other hand, whenever quantizers are used as a mean to reduce the quantity of information sent through a finite bandwidth channel, fast switchings traduce into an overly large number of transmissions per unit of time; [22]. To overcome this problem, in [22] a hysteretic quantizer has been proposed and analyzed in a consensus setting via quantized information. Part of our ongoing work consists of extending the analysis proposed in [22] to the general case of nonlinear systems as well as proposing some refinements of the model in [22], so as to ensure some robustness properties for the resulting model. First researches have shown that the general idea proposed in [82], consisting of capitalizing on input-to-state stability for the quantization free closed-loop system, can be successfully applied even to tackle this more involved problem. We would like to point out that such a quantizer is no longer a static nonlinearity but it is a hybrid dynamical systems. Hence, the tools we have been considering in this setting are the ones introduced in [56].

This page is intentionally left blank.

Appendices

APPENDIX A

SOME USEFUL RESULTS

Lemma A.1. Let $f\colon \mathbb{R} \to \mathbb{R}$ be a continuous function. Suppose that there exist $t, s \in \mathbb{R}$ with $s < t$ such that $f(s) = 0$, and $f(t) > 0$. Then, there exists $s' \geq s$ such that $f(s') = 0$ and $f(x) > 0$ for each $x \in (s', t]$

Proof. By continuity of f and the fact that $f(t) > 0$, there exists $\epsilon > 0$ such that $f(x) > 0$ for each $x \in [t - \epsilon, t]$. Define the set

$$\Omega \coloneqq \{\epsilon > 0 \colon \forall\, x \in [t - \epsilon, t]\ f(x) > 0\}.$$

Observe that Ω is non-empty, and furthermore $\Omega \subset [s, t]$. Define $\gamma = \sup \Omega$. Let $\{x_k\}$ be a sequence belonging to $(t - \gamma, t]$ for each $k \in \mathbb{N}$, and such that $\lim x_k = t - \gamma$. By continuity of f and the definition of the set Ω, it follows that $\lim f(x_k) = f(t - \gamma) \geq 0$. Now we prove that necessarily $f(t - \gamma) = 0$. By contradiction, assume that $f(t - \gamma) > 0$, then still by continuity of f, there exists $\gamma_2 > 0$ such that for each $x \in [t - \gamma - \gamma_2, t - \gamma]$, $f(x) > 0$. But this contradicts the fact that $\gamma = \sup \Omega$. Hence, setting $s' = t - \gamma$ establishes the result. ■

Lemma A.2. Let $X, Y \in \mathbb{R}^{n\times n}$ be two symmetric positive definite matrices and $U, V \in \mathbb{R}^n$, such that $UV^\mathsf{T} = \mathrm{I} - XY$. The following statements are equivalent

(a) $\det\left(\begin{bmatrix} Y & \mathrm{I} \\ \bullet & X \end{bmatrix}\right) \neq 0$

(b) $\det(VU) \neq 0$

(c) $\det\left(\begin{bmatrix} Y & V \\ \mathrm{I} & \mathbf{0} \end{bmatrix}\right) \neq 0$ and $\det(U) \neq 0$

Proof. First notice that (a) can be replaced with $\det(\mathrm{I} - XY) \neq 0$. Moreover, since by definition $\det(VU) = \det(\mathrm{I} - XY)$, (a) and (b) are equivalent. Now, we show that (c) and (b) are equivalent. To this end, observe that $\det\left(\begin{bmatrix} Y & V \\ \mathrm{I} & \mathbf{0} \end{bmatrix}\right) = (-1)^n \det(V)$, thus (c) and (b) are equivalent, concluding the proof. ■

Part II

State Estimation and Observer-based Control in the Presence of Sporadic Measurements

INTRODUCTION

General Overview

Recent technological advances have enabled the control of dynamical systems using data transmitted over communication networks or using digital devices. In these context, data can get lost or can only be available intermittently [62, 65, 129]. As an example, when the controller and the system to control are connected through a network, and an estimate of the plant state is needed, the classical paradigms of accessing the output of the plant continuously [88] do not apply and new approaches are required. This practical needed has brought to life a new research area aimed at developing observer schemes accounting the discrete nature of the available measurements; see, *e.g*, [1, 4, 6, 74, 92], just to cite a few. In these works, by assuming a periodical availability of the measured output, the authors propose a discrete-time approach to the state estimation problem. Such an approach consists of two stages. First the continuous-time plant is discretized, then a discrete-time observer is proposed to estimate the state of the discretized version of the plant. However, this approach entails three main drawbacks. The first drawback stems from the fact that the intersample is completely lost only studying the evolution of the estimation error at sampling times. In fact, with such a discrete-time approach, no explicit bounds on the estimation error in between consecutive samples are available. The second drawback is that any mismatch between the actual sampling time and that one used to discretize the plant model induces an error in the discrete-time description of the state estimation problem. The third drawback is that in many modern applications, such as networked control systems, the output of the plant is often accessible only sporadically, making the fundamental assumption of periodically measuring unrealistic; see, *e.g*, [62, 65, 129].

To address these issues, several strategies are presented in the literature. Such strategies essentially belong to two main families. The first one pertains to observers whose state is entirely reset, according to a suitable law, whenever a new measurement is available, and that run in open loop in between such events (continuous-discrete observers). This approach is, for instance, pursued in [3, 94]. The second family of strategies considers instead continuous-time observers, for which the output injection error in between consecutive samples is estimated via a continuous-time processing of the last received measurement. This approach is pursed, *e.g.*, in [73, 90, 101, 104]. However, we would like to point out that, except for the zero order sample-and-hold scheme in [90, 104], the design of observers within this family is essentially performed by an emulation approach. Such an approach consists of first designing a continuous-time observer, and then to evaluate the maximum allowable sampling period (MASP) the designed observer can withstand. On the other hand, in real applications, most of the time the design of the observer needs to be performed ensuring convergence of the estimation error for a given maximum sampling time. In other words, an effective design strategy should allow to consider the maximum allowable sampling interval as a design parameter.

The main aspect shared by the two families of observers illustrated here above is that the resulting observers exhibit both continuous-time and impulsive behaviors. Roughly speaking, the fact of having intermittent incoming measurements gives rise to observation schemes that need to instantaneously adapt their working principle according to the data streams. This fact of relying on observation schemes that experience continuous-time and impulsive behaviors foster to analyze such a schemes via the tools arising from the literature of hybrid dynamical systems. In particular, recently a comprehensive and solid framework for the analysis of hybrid dynamical systems has been presented in [56]. Although the modeling framework in [56] is solid and allows to deal with general hybrid dynamics, to the best of our knowledge, the design of observers in the presence of sporadic measurements via the tools in [56] has not received attention by the existing literature.

Another appealing aspect consists of analyzing the impact of sporadic measurement streams on observer-based controller architectures. Indeed, often the estimate provided by asymptotic observers is exploited to replace the actual plant state into static state feedback controller schemes; [128]. In the context of modern control systems, several settings can be considered. On the one hand, one can assume that, although the plant output is measurable sporadically, the plant input can be accessed at any time. This situation may occur, *e.g.*, when the output is measured via digital sensors with a low and time-varying sampling rate, or in distributed control systems, whenever the controller and the plant are co-located and plant measurements are sent to the controller via a data network; see, *e.g.*, [129]. On the other hand, in some real applications, temporal limitations can even affect the access to the plant input. As an example, in distributed systems, where the controller and the plant are located in different areas, the communication between the two systems happens via a shared channel handled by a supervisor. Such a supervisor alternatively allocates communication resources to the controller, to send control inputs toward the plant, and to the plant, to send measurements toward the controller; see [62]. Still within a distributed control systems framework, intermittent access to the plant input can be entailed also by package dropouts; see, *e.g.*, [112]. Another interesting case in which technological constraints involve intermittent actuation pertains to the case of low-rate actuators considered in [91].

Thus, in all these settings, the classical assumption considered in the literature of sampled-data systems consisting of assuming the sample and hold operations, of the measured output and of the control input, occur synchronously is overly restrictive.

An attempt to overcome this assumption is proposed, *e.g.*, in [50], where the authors, by pursuing a time-delay approach, propose a design strategy for an output feedback controller, guaranteeing an $\mathcal{H}_\infty$ performance, in the presence of aperiodic and asynchronous sampling and holding operations. Another work following a similar approach, though for the case of static state feedback controller is also presented in [91]. However, the proposed approach therein is to some extent intrinsically conservative due to the coarseness introduced by modeling the sampling and holding operations as processes introducing time-varying time delays.

Contribution

The contribution offered within this part of this dissertation aims at showing how the general hybrid systems framework proposed in [56] can be successfully adopted to model and design asymptotic observers for continuous-time LTI systems in the presence of intermittent measurements. In particular, we shall consider two observation schemes: The first one falls within the family of continuous-discrete observers considered in [3, 94], while the other falls within the family of observers considered in [73, 101, 104]. In addition, building on the first observation scheme, an observer-based controller architecture is proposed with the aim of stabilizing a continuous-time LTI system in the presence of both sporadic output measurements and input access. For such schemes computationally tractable design procedures will be illustrated and thoroughly discussed.

The contribution of the work presented in this part is twofold. On the one hand, resting on the general hybrid systems framework in [56] allows to come up with some completely novel observation schemes, whose design appear hardly tractable from a numerical standpoint by following alternative approaches as, *e.g.*, the one in [73]. On the other hand, adopting the general modeling framework in [56] allows to extend the derived results to deal with more involved problems of practical interest. For instance, the construction of the above mentioned observer-based controller essentially has the role to emphasize the flexibility and the modularity offered by the modeling framework in [56]. Other extensions are currently under preparation and will not presented in this thesis.

The remainder of this dissertation is organized as follows:

- Chapter 4 provides some general notions on hybrid systems as presented in [56].
- Chapter 5 illustrates the modeling and the design of a measurement triggered-jumps observer to exponentially estimate the state of a continuous-time LTI systems in the

presence of sporadic measurements. The results illustrated in this chapter are presented in [42, 44].

- Chapter 6 illustrates the modeling and the design of an observer with continuous intersample injection, still to exponentially estimate the state of a continuous-time LTI systems in the presence of sporadic measurements. Preliminaries results concerning the aspects illustrated in this chapter are presented in [41].
- Chapter 7 illustrates how the observer presented in Chapter 5 can be used to asymptotically stabilize a continuous-time LTI system in the presence of both sporadic measurements and intermittent input access. First results on this line of research can be found in [43].

Simulations of the hybrid systems contained in this part have been performed via the Hybrid Equations (HyEq) Toolbox [108].

4

PRELIMINARIES ON HYBRID SYSTEMS

"Beauty is the first test: there is no permanent place in the world for ugly mathematics."

– G. H. Hardy

4.1 Introduction

In this part of this dissertation, we rest on the hybrid system framework proposed in [56]. For this reason, within this chapter, we provide the main ingredients and the main definitions concerning hybrid systems. Notice that the list of notions given in this chapter is not an exhaustive one. In particular, for the sake of clarity, most of the definitions are given throughout the remainder of the dissertation. The aim of this chapter is to provide only the basic concepts and definitions needed to follow the results presented in the sequel of this dissertation. Thus, for a complete study of hybrid dynamical systems, the reader is referred to [56].

4.2 Hybrid systems: Modeling Framework and Basic Notions

In this part of this dissertation, we adopt the hybrid system framework proposed in [56]. In particular, we consider hybrid systems in the following form

$$\begin{cases} x \in C & \dot{x} \in F(x) \\ x \in D & x^+ \in G(x) \end{cases} \tag{4.1}$$

x is the state of the hybrid system, $\dot{x}$ stands for its velocity and x^+ indicates the value of the state after an instantaneous change. C indicates the set where the continuous evolution (flow) of the system can take place. Such an evolution is determined by the differential inclusion $\dot{x} \in F(x)$. D is the set wherein discrete evolution (jumps) can take place. Such an evolution is determined by the difference inclusion $x^+ \in G(x)$. In the sequel, according to [56], we name the objects defining the general hybrid system (4.1) as follows

- C is the flow set
- D is the jump set
- F is the flow map
- G is the jump map.

In particular, the four data (C, F, D, G) univocally define a hybrid system as in (4.1). For this reason, we refer to the four data (C, F, D, G) as data of the hybrid system (4.1). Specifically, the shorthand notation $\mathcal{H} = (C, F, D, G)$ stands for the hybrid system (4.1) represented by the data (C, F, D, G).

In this dissertation with focus on hybrid systems with state in $\mathbb{R}^n$. In that case, the data of the hybrid system (4.1) are defined precisely as follows:
Definition 4.1. The data of the hybrid system $\mathcal{H} = (C, F, D, G)$ with state in $\mathbb{R}^n$ are defined as follows.

- $C \subset \mathbb{R}^n$
- $F\colon \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ with $C \subset \operatorname{dom} F$
- $D \subset \mathbb{R}^n$
- $G\colon \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ with $D \subset \operatorname{dom} G$

4.3 Hybrid Time Domains and Solution Concept

In continuous-time systems, solutions are parameterized by a real scalar variable t, that is the time. Instead, in discrete-time systems, solutions are parameterized by an integer scalar variable $j \in \mathbb{N}$, that keeps track of the number of jumps or of the elapsed discrete steps. Since hybrid systems exhibit both continuous-time and discrete-time behaviors, it seems natural to parametrize solutions by means of two variables. The first one, $t \in \mathbb{R}_{\geq 0}$, representing the amount of time elapsed. The second one, $j \in \mathbb{N}$, keeps track of the number of the occurred jumps. However, a set $E \subset \mathbb{R}_{\geq 0} \times \mathbb{N}$ needs to satisfy some specific properties to provide the parametrization of a solution to some hybrid system. Such a properties are captured by the notion of hybrid time domain given next.

Definition 4.2 (Hybrid time domain)**.** A subset $E \subset \mathbb{R}_{\geq 0} \times \mathbb{N}$ is a compact hybrid time domain if

$$\bigcup_{j=0}^{J-1}([t_j, t_{j+1}], j)$$

for some finite sequences of times $0 = t_0 \leq t_1 \leq t_2 \leq \cdots \leq t_J$. It is a hybrid time domain if for all $(T, J) \in E$, $E \cap ([0, T] \times \{0, 1, \ldots, J\})$ is a compact hybrid time domain.

In the sequel, given a hybrid time domain E and $(t, j), (s, k) \in E$, the writing $(t, j) \succeq (s, k)$ means $t + j \geq s + k$. Furthermore, we indicate

$$\begin{aligned}
\sup\nolimits_t E &= \sup\{t \in \mathbb{R}_{\geq 0} \colon \exists j \in \mathbb{N} \,\text{such that}\, (t, j) \in E\} \\
\sup\nolimits_j E &= \sup\{j \in \mathbb{N} \colon \exists t \in \mathbb{R}_{\geq 0} \,\text{such that}\, (t, j) \in E\}.
\end{aligned}$$

Definition 4.3 (Hybrid arc)**.** A function $\phi \colon E \to \mathbb{R}^n$ is a hybrid arc if E is a hybrid time domain and if for each $j \in \mathbb{N}$, the function $t \mapsto \phi(t, j)$ is locally absolutely continuous on the interval $I^j = t \colon (t, j) \in E$.

Notice that from some j, the intervals I^j can be empty or being singleton. In such cases, the above requirement on absolutely continuity is not relevant. Here below, we provide a first categorization of hybrid arcs based on their properties. In particular, here below we list only the properties that are relevant within this dissertation, for an exhaustive categorization of hybrid arcs, the reader is refereed to [56].

Definition 4.4 (Types of hybrid arc)**.** A hybrid arc ϕ is called.

- nontrivial if $\operatorname{dom} \phi$ contains at least two points
- complete if $\operatorname{dom} \phi$ is unbounded
- Zeno if it is complete and $\sup_t \operatorname{dom} \phi < \infty$

Now we are in position to provide the following definition proving the concept of solution to hybrid systems used throughout the sequel of this dissertation.

Definition 4.5 (Solution to a hybrid system)**.** Given a hybrid system $\mathcal{H} = (C, F, D, G)$. A hybrid arc ϕ is a solution to $\mathcal{H}$ if $\phi(0, 0) \in \overline{C} \cup D$, and

(S1) for all $j \in \mathbb{N}_0$ such that $I^j := \{t \colon (t, j) \in \operatorname{dom} \phi\}$ has nonempty interior.

$$\begin{aligned}
\phi(t, j) &\in C & \forall t &\in \operatorname{Int} I^j, \\
\dot{\phi}(t, j) &\in F(\phi(t, j)) & \forall t &\in \operatorname{Int} I^j;
\end{aligned}$$

(S2) for all $(t, j) \in \operatorname{dom} \phi$ such that $(t, j+1) \in \operatorname{dom} \phi$,

$$\begin{aligned} &\phi(t, j) \in D, \\ &\phi(t, j+1) \in G(\phi(t, j)) \end{aligned}$$

Remark 4.1. Notice that, given a hybrid system, it is inappropriate to first select a hybrid time domain and then try finding a solution to the given hybrid system having the selected domain. In other words, it is the solution itself that determines its own domain.

Now, we provide another definition that extends the concept of maximal solution from continuous-time and discrete-time systems to hybrid systems.
Definition 4.6 (Maximal solutions)**.** A solution ϕ to $\mathcal{H}$ is maximal if there does not exist another solution ψ to $\mathcal{H}$ such that $\operatorname{dom} \phi \subset \operatorname{dom} \psi$ and $\phi(t, j) = \psi(t, j)$ for all $(t, j) \in \operatorname{dom} \phi$.

In the sequel of this dissertation, given a hybrid system $\mathcal{H}$, and a set S, $\mathcal{S}_{\mathcal{H}}(S)$ denotes the set of all maximal solution to $\mathcal{H}$ such that $\phi(0, 0) \in S$. If no set S is mentioned, then $\mathcal{S}_{\mathcal{H}}$ stands for the set of all maximal solutions to $\mathcal{H}$.

4.4 Basic Assumptions on Data

Before ending this chapter, let us consider the following assumption
Assumption 4.1 (Hybrid basic conditions)**.**

(A1) C and D are closed subsets of $\mathbb{R}^n$

(A2) $F\colon \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ is outer semicontinuous and locally bounded relative to C, $C \subset \operatorname{dom} F$, and $F(x)$ is convex valued for every $x \in C$

(A3) $G\colon \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ is outer semicontinuous and locally bounded relative to D, and $D \subset \operatorname{dom} G$

Such an assumption ensures that the considered hybrid system is well-posed in the sense specified in [56, Definition 6.2]; see [56, Theorem 6.8]. Well-posedness is a key property that is required for the applicability of a large number of results presented in [56]. We invite the reader to see [56] for further details on well-posed hybrid systems.

5

AN OBSERVER WITH MEASUREMENT-TRIGGERED JUMPS

"Success depends upon previous preparation, and without such preparation there is sure to be failure."

– Confucius

5.1 Introduction

THIS chapter deals with the state estimation problem for linear time-invariant (LTI) systems for which measurements of the output are available sporadically. To solve the considered problem, we provide an observer with jumps triggered by incoming measurements, which is studied in a hybrid systems framework. Specifically, the resulting system is written in estimation error coordinates and augmented with a timer variable that triggers the event of new measurements arriving. Then, the observer is performed to achieve global exponential stability (GES) of a closed set including the points for which the state of the plant and its estimate coincide. Furthermore, a computationally tractable procedure for the proposed observer is presented. Finally, the effectiveness of the proposed methodology is demonstrated in two numerical examples. The results presented in this chapter can be found in [44, 42].

5.2 Problem statement

5.2.1 System description

We consider continuous-time linear time-invariant systems of the form

$$\begin{aligned} \dot{z} &= Az + Bu \\ y &= Mz \end{aligned} \tag{5.1}$$

where $z \in \mathbb{R}^n$, $y \in \mathbb{R}^q$, and $u \in \mathbb{R}^p$ are, respectively, the state, the measured output, and the input of the system, while A, B and M are constant matrices of appropriate dimensions. We assume that the input u belongs to the class of measurable and locally bounded functions $u\colon [0, \infty) \to \mathbb{R}^p$. Our goal is to design an observer providing an estimate $\hat{z}$ of the state z when the output y is available only at some time instances t_k, $k \in \mathbb{N}$, not known *a priori* (a similar setup is considered in [100]).

We assume that the sequence $\{t_k\}_{k=1}^{\infty}$ is strictly increasing and unbounded, and that for such a sequence there exist two positive real scalars $T_1 \leq T_2$ such that

$$\begin{aligned} 0 \leq t_1 \leq T_2& \\ T_1 \leq t_{k+1} - t_k \leq T_2& \qquad \forall k \in \mathbb{N}. \end{aligned} \tag{5.2}$$

As also pointed out in [64], the lower bound in condition (5.2) prevents the existence of accumulation points in the sequence $\{t_k\}_{k=1}^{\infty}$, and, hence, avoids the existence of Zeno behaviors, which are typically undesired in practice. In fact, T_1 defines a strictly positive minimum time in between two consecutive incoming measurements. Furthermore, T_2 defines maximum time in between two consecutive incoming measurements. For this reason, we will refer to T_2 in the sequel as maximum sampling interval.

Since the information on the output y is available in an impulsive fashion, assuming that the arrival of a new measurement can be instantaneously detected, motivated by [3, 103], to solve the considered estimation problem, we consider an observer with jumps in its state following the law

$$\begin{cases} \dot{\hat{z}}(t) = A\hat{z}(t) + Bu(t) & \forall t \neq t_k, k \in \mathbb{N} \\ \hat{z}(t^+) = \hat{z}(t) + L(y(t) - M\hat{z}(t)) & \forall t = t_k, k \in \mathbb{N} \end{cases} \tag{5.3}$$

where L is a real matrix of appropriate dimensions to be designed. Note that, in between events, the observer runs in "open-loop" in the sense that no information of the output is used.

Remark 5.1. Assuming the knowledge of the input is not overly restrictive. Indeed, in many practical settings, all of the devices employed to control and supervise the plant may be embedded into the same system. Notice also that, often, the estimated state is part of a feedback controller (*e.g.* in linear observer-based controller architectures), in which case the input u is a static function of the estimated state that is perfectly known.

Along the lines of [109], the state estimation problem is formulated as a set stabilization problem. Namely, our goal is to design the matrix L such that the set wherein the plant state z and its estimate $\hat{z}$ coincide is globally exponentially stable for the plant (5.1) interconnected with the observer in (5.3). At this stage, as usual in estimation problems, we define the estimation error as

$$\varepsilon \coloneqq z - \hat{z}. \tag{5.4}$$

Thus, since at times t_k the plant state in unchanged, the error dynamics are given by the following dynamical system with jumps:

$$\begin{cases} \dot{\varepsilon}(t) = A\varepsilon(t) & \forall t \neq t_k, k \in \mathbb{N} \\ \varepsilon(t^+) = (\mathrm{I} - LM)\varepsilon(t) & \forall t = t_k, k \in \mathbb{N}. \end{cases} \tag{5.5}$$

Due to the linearity of system (5.1), the estimation error dynamics and the dynamics of z are decoupled. Then, for the purpose of estimation, one can effectively only consider system (5.5).

5.2.2 Hybrid Modeling

The fact that the observer experiences jumps when a new measurement is available and evolves according to a differential equation in between updates suggests that the updating process of the error dynamics can be described via a hybrid system. Due to this, we represent the whole system composed by the plant (5.1), the observer (5.3), and the logic triggering jumps as a hybrid system (see [81] where a similar approach is adopted to model a finite-time convergent observer).

The proposed hybrid systems approach requires to model the hidden time-driven mechanism triggering the jumps of the observer. To this end, in this work, and in a similar manner as in [19], we augment the state of the system with an auxiliary timer variable τ that keeps track of the duration of flows and triggers a jump whenever a certain condition is verified. This additional state allows to describe the time-driven triggering mechanism as a state-driven triggering mechanism, which leads to a model that can be efficiently represented by relying on the framework for hybrid systems proposed in [56]. More precisely, we make τ to decrease as ordinary time t increases and, whenever $\tau = 0$, reset it to any point in $[T_1, T_2]$, so as to enforce (5.2). After each jump, we require the system to flow again. The whole system composed by the estimation error ε and the timer variable τ can be represented by the following hybrid system, which we denote $\mathcal{H}_\varepsilon$

$$\mathcal{H}_\varepsilon \begin{cases} \left.\begin{array}{rcl} \dot{\varepsilon} & = & A\varepsilon \\ \dot{\tau} & = & -1 \end{array}\right\} & (\varepsilon,\tau) \in C \\ \\ \left.\begin{array}{rcl} \varepsilon^+ & = & (\mathrm{I} - LM)\varepsilon \\ \tau^+ & \in & [T_1, T_2] \end{array}\right\} & (\varepsilon,\tau) \in D \end{cases} \tag{5.6a}$$

where the flow set and the jump set are defined as

$$\begin{aligned} C &= \{(\varepsilon,\tau) \in \mathbb{R}^n \times \mathbb{R}_{\geq 0} \colon \tau \in [0, T_2]\} \\ D &= \{(\varepsilon,\tau) \in \mathbb{R}^n \times \mathbb{R}_{\geq 0} \colon \tau = 0\}\,. \end{aligned} \tag{5.6b}$$

The set-valued jump map allows to capture all possible measurement events within T_1 or T_2 units of time. Specifically, the hybrid model in (5.6) is able to characterize not only the behavior of the analyzed system for a given sequence $\{t_k\}_{k=1}^{\infty}$, but for any sequence satisfying (5.2). We denote the state of $\mathcal{H}_\varepsilon$ by

$$x = (\varepsilon, \tau)$$

and by f and G, respectively, the flow map and the jump map, *i.e.*,

$$f(x) = \begin{bmatrix} A\varepsilon \\ -1 \end{bmatrix} \qquad \forall x \in C \tag{5.7a}$$

$$G(x) = \begin{bmatrix} (\mathrm{I} - LM)\varepsilon \\ [T_1, T_2] \end{bmatrix} \qquad \forall x \in D. \tag{5.7b}$$

Remark 5.2. It is worthwhile to notice that the hybrid system $\mathcal{H}_\varepsilon$ satisfies Assumption 4.1. This assertion can be straightforwardly verified by inspection of the data of $\mathcal{H}_\varepsilon$. On the one hand, this property not only guarantees that the stability property exhibited for $\mathcal{H}_\varepsilon$ are somehow robust with respect to perturbations. However, in this dissertation we do not focus on perturbed hybrid systems and we refer to [56] for a complete treatment of this aspect. On the other hand, having Assumption 4.1 satisfied will be a crucial aspect in the sequel of this dissertation, being required for the derivation of some results.

Remark 5.3. To make the hybrid system (5.6) an accurate description of the real time-triggered phenomenon, which governs the feedback update process, the variable τ needs to belong to the interval $[0, T_2]$, property that is guaranteed by the definition of C and D.

In this chapter, we consider the following notion of global exponential stability (*GES*) of closed sets for a general hybrid system $\mathcal{H}$ in $\mathbb{R}^\ell$.

Definition 5.1. (GES [123]) Let $\mathcal{A} \subset \mathbb{R}^\ell$ be closed. The set $\mathcal{A}$ is said to be *globally exponentially stable* (GES) for the hybrid system $\mathcal{H}$ if there exist strictly positive real numbers λ, κ such that every solution ϕ to $\mathcal{H}$ satisfies for all $(t,j) \in \operatorname{dom}\phi$

$$|\phi(t,j)|_{\mathcal{A}} \leq \kappa e^{-\lambda(t+j)} |\phi(0,0)|_{\mathcal{A}}. \tag{5.8}$$

Then, by introducing the set[1]

$$\mathcal{A} = \{(\varepsilon, \tau) \in \mathbb{R}^n \times \mathbb{R}_{\geq 0} \colon \varepsilon = 0, \tau \in [0, T_2]\}\,. \tag{5.9}$$

the problem to solve is formulated as follows:

Problem 5.1. Given the matrices A, B, and M of appropriate dimensions and two positive scalars $T_1 \leq T_2$, design a matrix $L \in \mathbb{R}^{n\times q}$ such that the set $\mathcal{A}$ defined in (5.9) is GES for the hybrid system (5.6).

Remark 5.4. Concerning the existence of solutions to system (5.6), by relying on the concept of solution proposed in Definition 4.5, it is straightforward to check that for every initial condition $\phi(0,0) \in C \cup D$ there exists at least a nontrivial solution to (5.6) and that every maximal solution to (5.6) is complete. Notice that, although Definition 5.1 does not insist on completeness of maximal solutions, since the completeness requirement stated in Problem 5.1 is automatically satisfied by (5.6), solving Problem 5.1 ensures that the estimation error converges exponentially to zero as $t + j$ goes to infinity.

In addition, we can characterize the domain of the solutions to (5.6). Indeed, the variable τ, acting as a timer, guarantees that for every initial condition $\phi(0,0) \in C \cup D$, the domain of every maximal solution ϕ to (5.6) can be written as follows:

$$\operatorname{dom}\phi = \bigcup_{j\in\mathbb{N}_0} ([t_j, t_{j+1}]) \times \{j\} \tag{5.10a}$$

with

$$\begin{aligned} T_1 \leq t_{j+1} - t_j \ &\leq T_2 \quad \forall j \in \mathbb{N}_0 \setminus \{0\} \\ 0 \leq t_1 - t_0 \ &\leq T_2. \end{aligned} \tag{5.10b}$$

Furthermore, assuming $t_0 = 0$, the structure of the above hybrid time domain implies that for each $(t, j) \in \operatorname{dom}\phi$ we have

$$t \leq T_2(j+1) \tag{5.11}$$

the latter relation will play a key role in establishing GES of the set $\mathcal{A}$ for hybrid system (5.6).

5.3 Main Results

5.3.1 Conditions for GES

The following result provides conditions for GES of the set $\mathcal{A}$ defined in (5.9) for hybrid ystem (5.6).

Theorem 5.1. If there exist $P \in \mathcal{S}^n_+$, and a matrix $L \in \mathbb{R}^{n\times q}$ such that

$$(\mathrm{I} - LM)^\mathsf{T} e^{A^\mathsf{T} v} P e^{Av} (\mathrm{I} - LM) - P < \mathbf{0} \qquad \forall v \in [T_1, T_2] \tag{5.12}$$

[1]By the definition of system (5.6) and of the set $\mathcal{A}$, for every $x \in C \cup D \cup G(D)$, $|x|_{\mathcal{A}} = \|\varepsilon\|$.

then the set $\mathcal{A}$ defined in (5.9) is GES for the hybrid system (5.6).

The proof of the above theorem relies on the following Lemma.

Lemma 5.1. Let ϑ be a strictly negative real number. Pick

$$\gamma \in \left(0, \frac{|\vartheta|}{1+T_2}\right], R \in \left[\frac{|\vartheta| T_2}{1+T_2}, +\infty\right). \tag{5.13}$$

Let ϕ be any solution to the hybrid system (5.6). Then, for every $(t,j) \in \operatorname{dom}\phi$, one has

$$\vartheta j \leq R - \gamma(t+j). \tag{5.14}$$

Proof. From (5.14), by rearranging the terms, one gets

$$\gamma t + (\gamma + \vartheta) j - R \leq 0 \qquad \forall (t,j) \in \operatorname{dom}\phi \tag{5.15}$$

Now, pick any solution ϕ to (5.6). Now recall that from (5.11) for every $(t,j) \in \operatorname{dom}\phi$ one has

$$t \leq T_2(j+1) \tag{5.16}$$

then, for every strictly positive scalar γ, from the latter expression, one gets

$$\gamma t \leq \gamma T_2 j + \gamma T_2 \qquad \forall (t,j) \in \operatorname{dom}\phi. \tag{5.17}$$

Thus, by the virtue of the above bound, it turns out that (5.15) holds if

$$(\gamma T_2 + \gamma + \vartheta) j - R + \gamma T_2 \leq 0 \qquad \forall j \in \mathbb{N}_0 \tag{5.18}$$

which holds due to the selections considered in (5.13) for γ and R, concluding the proof. ■

Now we are in position to state the proof of Theorem 5.1

Proof. Consider the following Lyapunov function candidate for the hybrid system (5.6) defined for every $x \in \mathbb{R}^n \times \mathbb{R}_{\geq 0}$ and every $P \in \mathcal{S}_+^n$:

$$V(x) = \varepsilon^\mathsf{T} e^{A^\mathsf{T}\tau} P e^{A\tau} \varepsilon. \tag{5.19}$$

Note that there exist two positive scalars α_1, α_2 such that

$$\alpha_1 |x|_{\mathcal{A}}^2 \leq V(x) \leq \alpha_2 |x|_{\mathcal{A}}^2 \; \forall x \in C \cup D \cup G(D). \tag{5.20}$$

Specifically, due to the positive definiteness of P and the nonsingularity of the matrix $e^{A\tau}$ for every τ, by continuity arguments, one can set

$$\alpha_1 = \min_{\tau \in [0,T_2]} \lambda_{\min}\left(e^{A^\mathsf{T}\tau} P e^{A\tau}\right) \tag{5.21}$$

$$\alpha_2 = \max_{\tau \in [0,T_2]} \lambda_{\max}\left(e^{A^\mathsf{T}\tau} P e^{A\tau}\right) \tag{5.22}$$

where $\lambda_{\min}(\cdot)$ and $\lambda_{\max}(\cdot)$ denote, respectively, the smallest and the largest eigenvalue of their matrix argument. By straightforward calculations one gets

$$\nabla V(x) = \left(2e^{A^\mathsf{T}\tau}Pe^{A\tau}\varepsilon,\ \varepsilon^\mathsf{T}e^{A^\mathsf{T}\tau}(A^\mathsf{T}P + PA)e^{A\tau}\varepsilon\right).$$

Moreover, by exploiting the fact that the matrices $e^{A\tau}$ and A commute, one has

$$\langle \nabla V(x), f(x)\rangle = 0 \qquad \forall x \in C. \tag{5.23}$$

Now, observe that for every $g \in G(x)$, there exists a real scalar v belonging to the interval $[T_1, T_2]$ such that

$$g = \begin{bmatrix} (\mathrm{I} - LM)\varepsilon \\ v \end{bmatrix}.$$

Then, for every $g \in G(x)$, one has

$$\begin{aligned} V(g) - V(x) = {}& \varepsilon^\mathsf{T}(\mathrm{I} - LM)^\mathsf{T}e^{A^\mathsf{T}v}Pe^{Av}(\mathrm{I} - LM)\varepsilon \\ & - \varepsilon^\mathsf{T}e^{A^\mathsf{T}\tau}Pe^{A^\mathsf{T}\tau}\varepsilon. \end{aligned}$$

Furthermore, whenever $x \in D$, from (5.6b), we have that $\tau = 0$, which in turn implies

$$V(g) - V(x) = \varepsilon^\mathsf{T}\left((\mathrm{I} - LM)^\mathsf{T}e^{A^\mathsf{T}v}Pe^{Av}(\mathrm{I} - LM) - P\right)\varepsilon.$$

Hence, by virtue of relation (5.12), it follows that there exists a positive small enough scalar β such that, for every $x \in D, g \in G(x)$

$$V(g) - V(x) \leq -\beta\varepsilon^\mathsf{T}\varepsilon = -\beta|x|^2_{\mathcal{A}}. \tag{5.24}$$

Without loss of generality, assume that α_2 in (5.22) and β in (5.24) satisfy $1 - \frac{\beta}{\alpha_2} > 0$, which is always possible by picking β small enough. Define $\theta = \ln\left(1 - \frac{\beta}{\alpha_2}\right)$ and observe that $\theta < 0$. Then

$$V(g) \leq e^{\theta}V(x) \qquad \forall x \in D, g \in G(x). \tag{5.25}$$

Pick

$$\gamma \in \left(0, \frac{|\theta|}{1 + T_2}\right] \text{ and } R \in \left[\frac{T_2|\theta|}{1 + T_2}, \infty\right). \tag{5.26}$$

Let ϕ be a maximal solution to (5.6). As shown in the proof of [56, Proposition 3.29], thanks to (5.23) and (5.25), direct integration of $(t, j) \mapsto V(\phi(t, j))$ over $\operatorname{dom}\phi$ yields

$$V(\phi(t, j)) \leq e^{\theta j}V(\phi(0, 0)) \qquad \forall (t, j) \in \operatorname{dom}\phi. \tag{5.27}$$

Then, according to Lemma 5.1, due to the selection considered for γ and R in (5.26), from (5.27) one gets

$$\theta j \leq R - \gamma(t + j) \qquad \forall (t, j) \in \operatorname{dom}\phi \tag{5.28}$$

which, along with (5.20) and (5.27), leads to

$$|\phi(t,j)|_{\mathcal{A}} \leq e^{\frac{R}{2}}\sqrt{\frac{\alpha_2}{\alpha_1}}e^{-\frac{\gamma}{2}(t+j)}|\phi(0,0)|_{\mathcal{A}} \qquad \forall (t,j) \in \operatorname{dom}\phi. \tag{5.29}$$

Hence the set $\mathcal{A}$ defined in (5.9) is GES for system (5.6) concluding the proof. ■

Remark 5.5. Notice that assuming relation (5.12) to hold implies that the eigenvalues of $e^{Av}(\mathrm{I}-LM)$ are strictly contained in the unit circle for every v belonging to $[T_1,T_2]$. On the other hand, according to Sylvester's determinant theorem, $\operatorname{spec}(e^{Av}(\mathrm{I}-LM)) = \operatorname{spec}((\mathrm{I}-LM)e^{Av})$. Thus, the existence of a pair P,L satisfying condition (5.12) requires the detectability of the pair (e^{Av}, Me^{Av}) for each v belonging to $[T_1,T_2]$, which in turn, due to the nonsingularity of e^{Av} for any v and for any matrix A, is equivalent to the detectability of the pair (e^{Av}, M). Thus, it follows that Theorem 5.1 requires the sampled version of system (5.1) to be detectable for every v belonging to $[T_1,T_2]$, though this condition, in general, is only necessary. A similar remark is pointed out in [103].

5.3.2 Effect of Measurement Noise

So far, the measured output was assumed to be perfectly known at sampling times t_k, $k \in \mathbb{N}$. However, in a real-world setting, the measured output is affected by measurement noise. To quantify the robustness properties of our observer, denote the measurement noise as $\eta\colon \mathbb{R}_{\geq 0} \to \mathbb{R}^q$. Then, the measured output is

$$y = Mx + \eta.$$

This, in view of the definition of ε given in (5.4), suggests considering the following hybrid system with state $x = (\varepsilon,\tau) \in \mathbb{R}\times\mathbb{R}_{\geq 0}$ and input $\eta \in \mathbb{R}^q$

$$\mathcal{H}_\eta \begin{cases} \left.\begin{array}{rcl} \dot{\varepsilon} &=& A\varepsilon \\ \dot{\tau} &=& -1 \end{array}\right\} & (\varepsilon,\tau) \in C \\ \\ \left.\begin{array}{rcl} \varepsilon^+ &=& (\mathrm{I}-LM)\varepsilon - L\eta \\ \tau^+ &\in& [T_1,T_2] \end{array}\right\} & (\varepsilon,\tau) \in D. \end{cases} \tag{5.30}$$

For notational simplicity, in the sequel we use

$$\widetilde{G}(x,\eta) = \begin{bmatrix} (\mathrm{I}-LM)\varepsilon - L\eta \\ [T_1,T_2] \end{bmatrix}. \tag{5.31}$$

To study the effect of the measurement noise, we consider the input-to-state-stability (*ISS*) concept introduced in [115] for continuous-time nonlinear systems and extended to hybrid systems in [18]. Such a notion is given next for a general hybrid system $\mathcal{H}_d$ with state in $\mathbb{R}^\ell$, and input $d \in \mathbb{R}^s$. Before, consider the following notions of solution pair to $\mathcal{H}_d$, and the supremum norm for hybrid signals

Definition 5.2. Given an hybrid arc d, its superior norm at (t,j) is

$$\|d\|_{(t,j)} := \max\left\{ \underset{(s,k)\in \operatorname{dom} d\setminus\Gamma(d),(s,k)\preceq(t,j)}{\operatorname{ess.sup}} |d(s,k)| \quad , \quad \sup_{(s,k)\in\Gamma(d),(s,k)\preceq(t,j)} |d(s,k)| \right\}$$

where $\Gamma(d)$ denotes the set of all $(t,j) \in \operatorname{dom} d$ such that $(t,j+1) \in \operatorname{dom} d$; see [18] for further details.

Definition 5.3. A hybrid arc ϕ and a hybrid signal d is a solution pair (ϕ, d) to $\mathcal{H}_d = (F, G, C, D)$ if

- $\phi(0,0) \in \overline{C} \cup D$
- $\operatorname{dom}\phi = \operatorname{dom} d$
- for all $j \in \mathbb{N}$ and almost all t such that $(t,j) \in \operatorname{dom}\phi$

$$\phi(t,j) \in C, \quad \dot{\phi}(t,j) \in F(\phi(t,j), d(t,j))$$

- for all $(t,j) \in \operatorname{dom}\phi$ such that $(t,j+1) \in \operatorname{dom}\phi$

$$\phi(t,j) \in D, \quad \phi(t,j+1) \in G(\phi(t,j), d(t,j))$$

Building on these notions, let us consider the following definition.

Definition 5.4 ([18])**.** A hybrid system $\mathcal{H}_d$ is input-to-state-stable with respect to d and relatively to $\mathcal{A}$ if there exist $\gamma \in \mathcal{KL}$ and $\kappa \in \mathcal{K}$ such that each solution pair to $\mathcal{H}_d$ satisfies

$$|\phi(t,j)|_{\mathcal{A}} \leq \max\{\gamma(|\phi(0,0)|_{\mathcal{A}}, t+j), \kappa(\|d\|_{(t,j)})\} \tag{5.32}$$

for each $(t,j) \in \operatorname{dom}\phi$.

Remark 5.6. This extension of ISS to hybrid systems deals with hybrid signals as external perturbations. In our case, due to the continuous-time nature of the plant, the perturbation $t \mapsto \eta(t)$ acting on the measured output is a purely continuous-time signal. On the other hand, such a perturbation can be transformed into a hybrid signal to fit in the framework proposed by [18]. To this end, as shown in [105], given a solution ϕ to $\mathcal{H}_\eta$, the signal $t \mapsto \eta(t)$ can be represented as a hybrid signal $\eta_{\mathcal{H}}$ defined as

$$\eta_{\mathcal{H}}(t,j) := \eta(t) \qquad \forall (t,j) \in \operatorname{dom}\phi. \tag{5.33}$$

In particular $(\phi, \eta_{\mathcal{H}})$ is a solution pair to $\mathcal{H}_\eta$. Moreover, due to the form of $\eta_{\mathcal{H}}$, the hybrid sup norm $\|\eta_{\mathcal{H}}\|_{(t,j)}$ satisfies $\|\eta_{\mathcal{H}}\|_{(t,j)} = \|\eta\|_t$ for every $(t,j) \in \operatorname{dom}\phi$.

Remark 5.7. Notice that, since the Lyapunov function in (5.19) does not decrease during flows, the ISS Lyapunov condition for hybrid systems given in [18] cannot be employed in our setting. Thus, to show ISS of system (5.30) via the Lyapunov function given in Theorem 5.1, we couple strict decrease at jumps of such a function with the persistence of jumps enforced by the variable τ. This claim is formalized in the result given next.

Theorem 5.2. Let $T_1 \leq T_2$ be two positive real scalars. If there exist $P \in \mathcal{S}_+^n$, and a matrix $L \in \mathbb{R}^{n\times q}$ satisfying condition (5.12), then the hybrid system (5.30) is ISS with respect to η relatively to the set $\mathcal{A}$.

Proof. Consider the Lyapunov function defined in (5.19). Since the measurement noise η does not act on the flow map, as in the proof of Theorem 5.1, one gets

$$\langle \nabla V(x), f(x)\rangle = 0 \qquad \forall x \in C. \tag{5.34}$$

For any $(x,\eta) \in \mathbb{R} \times \mathbb{R}_{\geq 0} \times \mathbb{R}^q$ and for each $g \in \widetilde{G}(x,\eta)$ one gets

$$\begin{aligned} V(g) - V(x) =& \varepsilon^\mathsf{T}\Big((\mathrm{I}-LM)^\mathsf{T} e^{A^\mathsf{T}v} P e^{Av}(\mathrm{I}-LM) - e^{A^\mathsf{T}\tau} P e^{A\tau}\Big)\varepsilon - 2\eta^\mathsf{T} L^\mathsf{T} e^{A^\mathsf{T}v} P e^{Av}(\mathrm{I}-LM)\varepsilon \\ &+ \eta^\mathsf{T} L^\mathsf{T} e^{A^\mathsf{T}v} P e^{Av} L\eta \end{aligned}$$

where v is a real scalar belonging to the interval $[T_1, T_2]$. Whenever $x \in D$, from (5.6b), we have $\tau = 0$. Then, for each $x \in D$, $\eta \in \mathbb{R}^q$, $g \in \widetilde{G}(x,\eta)$, one gets

$$\begin{aligned} V(g) - V(x) =& \varepsilon^\mathsf{T}\left((\mathrm{I}-LM)^\mathsf{T} e^{A^\mathsf{T}v} P e^{Av}(\mathrm{I}-LM) - P\right)\varepsilon - 2\eta^\mathsf{T} L^\mathsf{T} e^{A^\mathsf{T}v} P e^{Av}(\mathrm{I}-LM)\varepsilon \\ &+ \eta^\mathsf{T} L^\mathsf{T} e^{A^\mathsf{T}v} P e^{Av} L\eta. \end{aligned} \tag{5.35}$$

Moreover, from (5.12), there exists a small enough positive real scalar β such that, for every $v \in [T_1, T_2]$ and every ε

$$\varepsilon^\mathsf{T}\Big((\mathrm{I}-LM)^\mathsf{T} e^{A^\mathsf{T}v} P e^{Av}(\mathrm{I}-LM) - P\Big)\varepsilon \leq -\beta\varepsilon^\mathsf{T}\varepsilon. \tag{5.36}$$

Now recall that for every $a, b \in \mathbb{R}^n$, $2a^\mathsf{T} b \leq \omega a^\mathsf{T} a + \omega^{-1} b^\mathsf{T} b$ for every positive real scalar ω. From (5.35) and (5.36), setting $a = \varepsilon$, $b^\mathsf{T} = -\eta^\mathsf{T} L^\mathsf{T} e^{A^\mathsf{T}v} P e^{Av}(\mathrm{I}-LM)$, and $\omega = \frac{\beta}{2}$ yields

$$V(g) - V(x) \leq -\frac{1}{2}\beta\varepsilon^\mathsf{T}\varepsilon + \eta^\mathsf{T}\eta\left\| L^\mathsf{T} e^{A^\mathsf{T}v} P\Big(\mathrm{I}\frac{2}{\beta} + e^{Av}(\mathrm{I}-LM)(\mathrm{I}-LM)^\mathsf{T} e^{A^\mathsf{T}v} P\Big) e^{Av} L\right\|. \tag{5.37}$$

Moreover, thanks to (5.12), one has $\|(\mathrm{I}-LM)^\mathsf{T} e^{A^\mathsf{T}v} P e^{Av}(\mathrm{I}-LM)\| < \|P\|$. Thus, from (5.37), it follows $V(g) - V(x) \leq -\frac{1}{2}\beta\varepsilon^\mathsf{T}\varepsilon + \rho\|L\|^2\eta^\mathsf{T}\eta$, where

$$\rho = \|P\|\left(\frac{2}{\beta} + \|P\|\right)\max_{v\in[T_1,T_2]}\Big(\|e^{A^\mathsf{T}v}\|^2\Big).$$

The above relationship, together with (5.20), yields

$$V(g) \leq e^\theta V(x) + \|L\|^2\rho\eta^\mathsf{T}\eta \quad \forall x \in D, \eta \in \mathbb{R}^q, g \in \widetilde{G}(x,\eta) \tag{5.38}$$

where $\theta = \ln\Big(1 - \frac{\beta}{2\alpha_2}\Big)$ and α_2 is defined in (5.22). Therefore, from (5.38) and (5.23), by considering the domain of the solutions to (5.30), which is given in (5.10), it turns out that given any maximal solution pair (ϕ, η) to (5.30), one gets

$$V(\phi(t,0)) = V(\phi(0,0)) \qquad \forall t \in [0, t_1] \tag{5.39a}$$

$$\begin{gathered} V(\phi(t,j)) \leq e^{\theta j} V(\phi(0,0)) + \rho \|L\|^2 \sum_{i=0}^{j-1} e^{\theta(j-1-i)} \|\eta(t_{i+1}, i+1)\|^2 \\ \forall (t,j) \in \operatorname{dom} \phi \qquad \text{with } j \geq 1 \end{gathered} \tag{5.39b}$$

Furthermore, with θ negative as in the proof of Theorem 5.1, for each $(t, j) \in \operatorname{dom} \phi$ such that $j \geq 1$, we have

$$V(\phi(t,j)) \leq e^{\theta j} V(\phi(0,0)) + \frac{\rho e^{-\theta} \|L\|^2}{e^{-\theta} - 1} \|\eta\|^2_{(t,j)}. \tag{5.40}$$

Since the input dependent term in the right-hand side of (5.40) is nonnegative, by combining it with (5.39a) and (5.40), we obtain for each $(t, j) \in \operatorname{dom} \phi$,

$$V(\phi(t,j)) \leq e^{\theta j} V(\phi(0,0)) + \frac{\rho e^{-\theta} \|L\|^2}{e^{-\theta} - 1} \|\eta\|^2_{(t,j)}, \tag{5.41}$$

further using (5.20) one gets

$$|\phi(t,j)|^2_{\mathcal{A}} \leq \frac{\alpha_2}{\alpha_1} e^{\theta j} |\phi(0,0)|^2_{\mathcal{A}} + \frac{\rho e^{-\theta} \|L\|^2}{(e^{-\theta} - 1)\alpha_1} \|\eta\|^2_{(t,j)}. \tag{5.42}$$

Now, by following the same arguments in the proof of Theorem 5.1, for some (solution independent) positive real scalars γ, R, from (5.42) one gets

$$|\phi(t,j)|^2_{\mathcal{A}} \leq e^{-\gamma(t+j)} e^{R} \frac{\alpha_2}{\alpha_1} |\phi(0,0)|^2_{\mathcal{A}} + \frac{\rho e^{-\theta} \|L\|^2}{(e^{-\theta} - 1)\alpha_1} \|\eta\|^2_{(t,j)} \tag{5.43}$$

or equivalently

$$|\phi(t,j)|_{\mathcal{A}} \leq \max \left\{ \sqrt{2 \frac{\alpha_2}{\alpha_1}} e^{\frac{R}{2}} e^{-\frac{\gamma(t+j)}{2}} |\phi(0,0)|_{\mathcal{A}}, \sqrt{\frac{2\rho e^{-\theta}}{(e^{-\theta} - 1)\alpha_1}} \|L\| \|\eta\|_{(t,j)} \right\} \tag{5.44}$$

Thus, according to Definition 5.4, the hybrid system (5.30) is ISS with respect to η (relatively to the set $\mathcal{A}$). ■

Remark 5.8. The above result allows to conclude that, in the considered case, condition (5.12) actually suffices to guarantee the ISS property for hybrid system (5.30), and there is no need in finding an ISS-Lyapunov function as defined in [18].

5.4 Observer Design

In the previous section a condition to guarantee global exponentially stability and input-to-state-stability, respectively, for systems (5.6) and (5.30) was provided. However, due to its form, such a condition is not computationally tractable to obtain a solution to Problem 5.1. Indeed, from a numerical standpoint, condition (5.12) has two drawbacks: it is not linear in P and L, and it needs to be verified for infinitely many values of v. The relevance of the second drawback is evident at a first sight, while the lack of linearity is a severe constraint, since the solution to nonlinear matrix inequalities often lead to NP-hard problems; see *e.g.*, [15]. Thus, to make the problem numerically tractable, further work is needed. To this end, the following result provides a first step toward an LMI-based design procedure for the proposed observer.

Proposition 5.1. *Let $P \in \mathcal{S}_+^n$ and $L \in \mathbb{R}^{n\times q}$. Then, (5.12) holds if*

$$\begin{bmatrix} -\operatorname{He}(F) & F - FLM & e^{A^\mathsf{T} v}P \\ \bullet & -P & \mathbf{0} \\ \bullet & \bullet & -P \end{bmatrix} < \mathbf{0} \qquad \forall v \in [T_1, T_2] \tag{5.45}$$

is feasible with respect to $F \in \mathbb{R}^{n\times n}$.

Proof. The proof carried out here is inspired by [99]. Specifically, set

$$Z = \begin{bmatrix} e^{Av}Pe^{A^\mathsf{T} v} & \mathbf{0} \\ \mathbf{0} & -P \end{bmatrix}, S = \begin{bmatrix} \mathrm{I} - LM \\ \mathrm{I} \end{bmatrix}, Y = \begin{bmatrix} \mathbf{0} \\ \mathrm{I} \end{bmatrix}.$$

Then, condition (5.12) can be rewritten as

$$S^\mathsf{T} Z S < \mathbf{0} \tag{5.46}$$

while the positive definiteness of P can be expressed equivalently by requiring that

$$Y^\mathsf{T} Z Y < \mathbf{0}. \tag{5.47}$$

Thus, by the projection lemma [52], (5.46) and (5.47) are satisfied if there exists a matrix F such that

$$\begin{bmatrix} e^{A^\mathsf{T} v}Pe^{Av} - \operatorname{He}(F) & F - FLM \\ \bullet & -P \end{bmatrix} < \mathbf{0}. \tag{5.48}$$

Moreover, by Schur complement, from (5.48) one gets

$$\begin{bmatrix} -\operatorname{He}(F) & F - FLM & e^{A^\mathsf{T} v} \\ \bullet & -P & \mathbf{0} \\ \bullet & \bullet & -P^{-1} \end{bmatrix} < \mathbf{0} \tag{5.49}$$

and finally, pre-and-post multiplying by $\operatorname{diag}(\mathrm{I}, \mathrm{I}, P)$ yields the left-hand side matrix in (5.45), concluding the proof. ■

Remark 5.9. Notice that by setting $FL = J$, condition (5.45) turns into a parametric LMI in v, with respect to the unknown matrices F, J, and P.

Proposition 5.1, along with the above remark, provides a sufficient condition to (5.12), which is linear in the decision variable F, J and P. Nevertheless, the obtained condition still has to be verified for infinitely many values of v. This situation is rather common in the literature of sampled data and impulsive systems; see, *e.g.*, [64] and the references therein. A general procedure to overcome this issue consists of embedding the term e^{Av}, with v in the interval $[T_1, T_2]$, into a polytope, (a convex set having a finite number of extreme points [125]). Namely, one needs to find some matrices $X_1, X_2, \dots, X_\nu \in \mathbb{R}^{n\times n}$ such that $e^{Av} \in \operatorname{co}\{X_1, X_2, \dots, X_\nu\}$ whenever $v \in [T_1, T_2]$. Throughout the sequel, we refer to such a polytope as polytopic overapproximation or polytopic embedding of e^{Av} on $[T_1, T_2]$. Then, by exploiting the linearity of condition (5.45) with respect to e^{Av}, one can obtain a finite set of LMIs, whose satisfaction implies (5.45) to hold. This approach is formalized for our case in the result given next.

Corollary 5.1. Let $X_1, X_2, \dots, X_\nu$ be matrices such that $e^{A[T_1,T_2]} \in \operatorname{co}\{X_1, X_2, \dots, X_\nu\}$. If there exist $P \in \mathcal{S}_+^n$, a matrix $J \in \mathbb{R}^{n\times q}$, and a matrix $F \in \mathbb{R}^{n\times n}$ such that, for every $i \in \{1, \dots, \nu\}$,

$$\begin{bmatrix} -\operatorname{He}(F) & F - JM & X_i^\mathsf{T} P \\ \bullet & -P & \mathbf{0} \\ \bullet & \bullet & -P \end{bmatrix} < \mathbf{0} \tag{5.50}$$

then the matrices P and $L = F^{-1}J$ satisfy condition (5.12).

Proof. Since $e^{Av} \in \operatorname{co}\{X_1, X_2, \dots, X_\nu\}$ whenever $v \in [T_1, T_2]$, then there exist non-negative functions $\xi_1, \xi_2, \dots, \xi_\nu$, such that for each $v \in [T_1, T_2]$

$$e^{Av} = \sum_{i=1}^{\nu} \xi_i(v) X_i, \qquad \sum_{i=1}^{\nu} \xi_i(v) = 1. \tag{5.51}$$

Then, replacing in left-hand side of (5.45) the term e^{Av} with the expression given in the left-hand side of (5.51) leads to

$$\begin{bmatrix} -\operatorname{He}(F) & F - JM & \sum_{i=1}^{\nu} \xi_i(v) X_i^\mathsf{T} P \\ \bullet & -P & \mathbf{0} \\ \bullet & \bullet & -P \end{bmatrix} \tag{5.52}$$

which, thanks to the constraint on each $\xi_i(v)$ given in the right-hand side of (5.51), can be equivalently rewritten as

$$\sum_{i=1}^{\nu} \xi_i(v) \begin{bmatrix} -\operatorname{He}(F) & F - JM & X_i^\mathsf{T} P \\ \bullet & -P & \mathbf{0} \\ \bullet & \bullet & -P \end{bmatrix} \tag{5.53}$$

Hence, by the virtue of (5.50) and Proposition 5.1 matrices P and $L = F^{-1}J$ satisfy condition (5.12) and this concludes the proof. ■

The previous result allows, once a polytopic embedding of the term e^{Av} is known, to design the proposed observer via the solution to a finite number of linear matrix inequalities. The next subsection illustrates a possible technique to build such an embedding.

5.4.1 Polytopic Embedding

The derivation of a polytopic overapproximation of the exponential matrix on a given compact interval is recognized in the literature as a difficult problem; see [29, 60]. In [60] an exhaustive comparison between several kinds of overapproximations is presented and the authors suggest that two classes of approaches can be pursued to determine polytopic overapproximations of the matrix exponential term on a given compact interval. In the sequel, for any interval $\mathcal{I} \subset \mathbb{R}$, we denote

$$e^{A\mathcal{I}} := \{Y \in \mathbb{R}^{n\times n} \colon \exists v \in \mathcal{I} \text{ such that } Y = e^{Av}\}.$$

The first approach aims at determining a finite number of matrices $F_1, F_2, \dots, F_\nu \in \mathbb{R}^{n\times n}$ such that $e^{A\mathcal{I}} \in \operatorname{co}\{F_1, F_2, \dots, F_\nu\}$ for a given compact interval $\mathcal{I}$. This approach is commonly called *without uncertainties*. The other approach leads to a finite numbers of matrices $F_1, F_2, \dots, F_\mu \in \mathbb{R}^{n\times n}$ and a norm bounded uncertainty $\Delta(v) \in \mathbb{R}^{n\times n}$ such that, for every v belonging to a given compact interval, $e^{Av} = \sum_{i=1}^{\mu} \alpha_i(v)F_i + \Delta(v)$ for some positive scalar functions $\alpha_1, \dots, \alpha_\mu$ with $\sum_{i=1}^{\mu} \alpha_i(v) = 1$. This approach is commonly called *with uncertainties*. On the one hand, the approaches with uncertainties allow, in general, to obtain tighter overapproximations than those without uncertainties; see [29]. On the other hand, managing bounded uncertainties to build a design procedure can be hard, although in [63] a possible two-stage design procedure is proposed to cope with this issue.

In this dissertation, we propose a novel methodology to build a polytopic embedding without uncertainties. Such a methodology is based on the well known expansion of the matrix exponential based on residue matrices. By arranging the eigenvalues of the matrix A in a way such that the first σ_r are real and distinct, the following σ_c are complex and distinct, and the remaining σ_c are the conjugates of the previous ones, such an expression is given by

$$\begin{aligned} e^{Av} &= \sum_{i=1}^{\sigma_r}\sum_{j=1}^{m_i^r} R_{ij}e^{\lambda_i v}\frac{v^{j-1}}{(j-1)!} \\ &+ \sum_{i=\sigma_r+1}^{\sigma_r+\sigma_c}\sum_{j=1}^{m_i^c} 2e^{\Re(\lambda_i)v}\Big(\Re(R_{ij})\cos(\Im(\lambda_i)v) - \Im(R_{ij})\sin(\Im(\lambda_i)v)\Big)\frac{v^{j-1}}{(j-1)!} \end{aligned} \tag{5.54}$$

The constants m_i^r and m_i^c are, respectively, the multiplicity of the real eigenvalue λ_i and of the complex-conjugate eigenvalue pair λ_i, λ_i^* in the minimal polynomial of the matrix A.

The matrices R_{ij} are real $n \times n$ matrices corresponding to the residues associated to the partial fraction expansion of the rational matrix $(s\mathrm{I} - A)^{-1}$. The advantage of the proposed method lies in the fact that there exist several methods to compute the residues matrices. For instance, in this work, we rely on the procedure proposed in [80].

Remark 5.10. Although the above expansion based on residue matrices concerns the multiplicity of each eigenvalues in the minimal polynomial of A, the knowledge of such a minimal polynomial is not required to start with the application of proposed methodology. Indeed, as a first step, one can assume without any loss of generality that the minimal polynomial of A coincides with its characteristic polynomial and compute the residues matrices for each eigenvalues according to its multiplicity in the characteristic polynomial. Then, for each eigenvalue λ, if the multiplicity of λ in the minimal polynomial of A is less than the one in the characteristic polynomial, higher order residues are automatically equal to zero. This feature is ensured in the algorithm proposed in [80]; see [80, Section 4 and Example 1]. Therefore, from a practical view point, as a first step, one can for each eigenvalue (depending on its multiplicity in the characteristic polynomial) compute all the related residues. Then, by neglecting the ones equal to zero (the selection of a certain threshold can be required in finite-precision implementations), one gets the right residues expansion.

Once the residue matrices are known, to build a polytopic embedding of e^{Av} one can proceed in a similar manner as in [29]. In particular, define for each $i = 1, 2, \dots, \sigma_r$ and for each $v \in [T_1, T_2]$

$$
\begin{aligned}
&\beta_i \colon v \mapsto \begin{bmatrix} \beta_{i1}(v) & \beta_{i2}(v) & \dots & \beta_{im_i^r}(v) \end{bmatrix} := \begin{bmatrix} e^{\lambda_i v} & e^{\lambda_i v} v & \dots & e^{\lambda_i v} \frac{v^{m_i^r - 1}}{(m_i^r - 1)!} \end{bmatrix} \\
&\widehat{R}_i := \begin{bmatrix} R_{i1} & R_{i2} & \dots & R_{im_i^r} \end{bmatrix}
\end{aligned}
$$

and set for each $v \in [T_1, T_2]$

$$
\begin{aligned}
\beta \colon v &\mapsto \begin{bmatrix} \beta_1(v) & \beta_2(v) & \dots & \beta_{\sigma_r}(v) \end{bmatrix}^\mathsf{T} \\
\Psi &:= \begin{bmatrix} \widehat{R}_1 & \widehat{R}_2 & \dots & \widehat{R}_{\sigma_r} \end{bmatrix}
\end{aligned}
$$

Define for each $i = \sigma_r + 1, \sigma_r + 2, \dots, \sigma_r + \sigma_c$ and for each $v \in [T_1, T_2]$

$$
\begin{aligned}
\gamma_i \colon v \mapsto &\begin{bmatrix} \gamma_{i1}(v) & \gamma_{i2}(v) & \dots & \gamma_{im_i^c}(v) \end{bmatrix} \\
:= &\begin{bmatrix} 2e^{\Re(\lambda_i)v} \cos(\Im(\lambda_i)v) & 2e^{\Re(\lambda_i)v} \cos(\Im(\lambda_i)v)v & \dots & 2e^{\Re(\lambda_i)v} \cos(\Im(\lambda_i)v) \frac{v^{m_i^c - 1}}{(m_i^c - 1)!} \end{bmatrix}
\end{aligned}
$$

$$
\begin{aligned}
\gamma_i'(v) \colon v \mapsto &\begin{bmatrix} \gamma_{i1}'(v) & \gamma_{i2}'(v) & \dots & \gamma_{im_i^c}'(v) \end{bmatrix} \\
:= &\begin{bmatrix} -2e^{\Re(\lambda_i)v} \sin(\Im(\lambda_i)v) & -2e^{\Re(\lambda_i)v} \sin(\Im(\lambda_i)v)v & \dots & -2e^{\Re(\lambda_i)v} \sin(\Im(\lambda_i)v) \frac{v^{m_i^c - 1}}{(m_i^c - 1)!} \end{bmatrix}
\end{aligned}
$$

$$
\widehat{Q}_i := \begin{bmatrix} \Re(R_{i1}) & \Re(R_{i2}) & \dots & \Re(R_{im^{c_i}}) \end{bmatrix}
$$

and set for each $v \in [T_1, T_2]$,

$$
\begin{aligned}
\gamma &\colon v \mapsto \begin{bmatrix} \gamma_{\sigma_r+1}(v) & \gamma_{\sigma_r+2}(v) & \dots & \gamma_{\sigma_r+1+\sigma_c}(v) \end{bmatrix}^{\mathsf{T}} \\
\gamma' &\colon v \mapsto \begin{bmatrix} \gamma'_{\sigma_r+1}(v) & \gamma'_{\sigma_r+2}(v) & \dots & \gamma_{\sigma'_r+1+\sigma_c}(v) \end{bmatrix}^{\mathsf{T}} \\
\Phi &\coloneqq \begin{bmatrix} \widehat{Q}_{\sigma_r+1} & \widehat{Q}_{\sigma_r+2} & \dots & \widehat{Q}_{\sigma_r+1+\sigma_c} \end{bmatrix}.
\end{aligned}
$$

The definition of the above quantities leads to the following equivalent writing for (5.54) for each $v \in [T_1, T_2]$

$$
e^{Av} = \begin{bmatrix} \Psi & \Phi & \Phi^\star \end{bmatrix} \left(\begin{bmatrix} \beta(v) \\ \gamma(v) \\ \gamma^\star(v) \end{bmatrix} \otimes \mathrm{I}_n \right). \tag{5.55}
$$

The above writing allows to make a separation between constant elements and functions of v appearing in (5.54), which is useful to build up a polytopic embedding for such an expression. To this aim, firstly observe that

$$
\operatorname{rge}(\beta \times \gamma \times \gamma') \subset \operatorname{rge}\beta \times \operatorname{rge}\gamma \times \operatorname{rge}\gamma'. \tag{5.56}
$$

Moreover, by defining the following quantities: for each $i \in \{1, 2, \dots, \sigma_r\}$

$$
\begin{aligned}
\overline{\beta_{ij}} &= \max_{v \in [T_1, T_2]} e^{\lambda_i v} \frac{v^{j-1}}{(j-1)!} \qquad j \in \{1, 2, \dots, m_i^r\} \\
\underline{\beta_{ij}} &= \min_{v \in [T_1, T_2]} e^{\lambda_i v} \frac{v^{j-1}}{(j-1)!} \qquad j \in \{1, 2, \dots, m_i^r\}
\end{aligned} \tag{5.57a}
$$

and for each $i \in \{\sigma_r + 1, \sigma_r + 2, \dots, \sigma_r + \sigma_c\}$

$$
\begin{aligned}
\overline{\gamma_{ij}} &= \max_{v \in [T_1, T_2]} 2e^{\Re(\lambda_i)v} \cos(\Im(\lambda_i)v) \frac{v^{j-1}}{(j-1)!} \qquad j \in \{1, 2, \dots, m_i^c\} \\
\underline{\gamma_{ij}} &= \min_{v \in [T_1, T_2]} 2e^{\Re(\lambda_i)v} \cos(\Im(\lambda_i)v) \frac{v^{j-1}}{(j-1)!} \qquad j \in \{1, 2, \dots, m_i^c\} \\
\overline{\gamma_{ij}^*} &= \max_{v \in [T_1, T_2]} -2e^{\Re(\lambda_i)v} \sin(\Im(\lambda_i)v) \frac{v^{j-1}}{(j-1)!} \qquad j \in \{1, 2, \dots, m_i^c\} \\
\underline{\gamma_{ij}^*} &= \min_{v \in [T_1, T_2]} -2e^{\Re(\lambda_i)v} \sin(\Im(\lambda_i)v) \frac{v^{j-1}}{(j-1)!} \qquad j \in \{1, 2, \dots, m_i^c\}
\end{aligned} \tag{5.57b}
$$

by continuity of the functions involved in (5.55), it turns out that[2]

$$\begin{aligned}
\operatorname{rge}\beta &\subset \bigtimes_{i=1}^{\sigma_r}\bigtimes_{j=1}^{m_i^r}\operatorname{rge}\beta_{ij} = \bigtimes_{i=1}^{\sigma_r}\bigtimes_{j=1}^{m_i^r}\operatorname{co}\{\overline{\beta_{ij}},\underline{\beta_{ij}}\} = \operatorname{co}\bigtimes_{i=1}^{\sigma_r}\bigtimes_{j=1}^{m_i^r}\{\overline{\beta_{ij}},\underline{\beta_{ij}}\} \coloneqq \operatorname{co}\Theta\\
\operatorname{rge}\gamma &\subset \bigtimes_{i=\sigma_r+1}^{\sigma_r+\sigma_c}\bigtimes_{j=1}^{m_i^c}\operatorname{rge}\gamma_{ij} = \bigtimes_{i=\sigma_r+1}^{\sigma_r+\sigma_c}\bigtimes_{j=1}^{m_i^c}\operatorname{co}\{\overline{\gamma_{ij}},\underline{\gamma_{ij}}\} = \operatorname{co}\bigtimes_{i=\sigma_r+1}^{\sigma_r+\sigma_c}\bigtimes_{j=1}^{m_i^c}\{\overline{\gamma_{ij}},\underline{\gamma_{ij}}\} \coloneqq \operatorname{co}\Gamma \qquad (5.58a)\\
\operatorname{rge}\gamma' &\subset \bigtimes_{i=\sigma_r+1}^{\sigma_r+\sigma_c}\bigtimes_{j=1}^{m_i^c}\operatorname{rge}\gamma'_{ij} = \bigtimes_{i=\sigma_r+1}^{\sigma_r+\sigma_c}\bigtimes_{j=1}^{m_i^c}\operatorname{co}\{\overline{\gamma'_{ij}},\underline{\gamma'_{ij}}\} = \operatorname{co}\bigtimes_{i=\sigma_r+1}^{\sigma_r+\sigma_c}\bigtimes_{j=1}^{m_i^c}\{\overline{\gamma'_{ij}},\underline{\gamma'_{ij}}\} \coloneqq \operatorname{co}\Gamma'
\end{aligned}$$

Thus from (5.55) and (5.56), via the above expressions, one gets

$$\begin{aligned}
e^{A[T_1,T_2]} &\subset \left\{\begin{bmatrix}\Psi & \Phi & \Phi^\star\end{bmatrix}(\xi\otimes \mathrm{I}_n) \mid \xi\in\operatorname{co}(\Theta\times\Gamma\times\Gamma')\right\}\\
&= \operatorname{co}\left(\underbrace{\begin{bmatrix}\Psi & \Phi & \Phi^\star\end{bmatrix}((\Theta\times\Gamma\times\Gamma')\otimes \mathrm{I}_n)}_{\Omega}\right).
\end{aligned} \qquad (5.58b)$$

Let us remark that the set Ω is a finite point set, hence each element belonging to its convex-hull is the convex combination of a finite number of elements in Ω. Specifically,

$$\nu = \mathbf{card}(\Omega) = \mathbf{card}(\Theta\times\Gamma\times\Gamma') = \mathbf{card}(\Theta)\mathbf{card}(\Gamma)\mathbf{card}(\Gamma') \le 2^{\sum_{i=1}^{\sigma_r} m_i^r}2^{2\sum_{i=\sigma_r+1}^{\sigma_r+\sigma_c} m_i^c} = 2^n.$$

Therefore, let $X_1, X_2, \dots, X_\nu$ be the matrices such that

$$\Omega = \{X_1, X_2, \dots, X_\nu\} \qquad (5.58c)$$

then for each $v\in[T_1,T_2]$,

$$e^{Av}\in\operatorname{co}\{X_1,X_2,\dots,X_\nu\}.$$

Remark 5.11. The most laborious part of the proposed technique, namely the computation of the residue matrices, does not depend on the considered interval $[T_1, T_2]$. Thus, for a given matrix A, once the residues are known and stored, the construction of the needed polytopic embedding only requires the computation of the extrema of a finite number of continuous scalars functions on a compact interval. Notice that although the proposed embedding technique could lead to similar results to the ones proposed in [29], our methodology does not require either the derivation of the real Jordan form of A or its minimal polynomial. Moreover, the proposed methodology is systematic and does not require dedicated strategies depending on the multiplicity of the eigenvalues.

Remark 5.12. As the tightness of the resulting polytopic embedding is not taken into account by the procedure itself, the resulting overapproximation can be rather conservative. However, although this conservatism plays a relevant role in analysis problems (where one is interested in obtaining a description of the exponential matrix as tight as possible), in

[2]Here we used the fact that given $S_1\subset\mathbb{R}^{n_1}, S_2\subset\mathbb{R}^{n_2}\dots, S_m\subset\mathbb{R}^{n_m}$ any sets, then $\operatorname{co}\bigtimes_{i=1}^m S_i = \bigtimes_{i=1}^s \operatorname{co} S_i$; see [13].

our case, being the final aim obtaining a design procedure, overapproximation tightness is not excessively crucial. Nevertheless, if needed, the overapproximating polytope can be made tighter by subdividing the interval $[T_1, T_2]$ in N subintervals and then by applying the proposed procedure on each subinterval. Specifically, the proposed technique operated on every subinterval leads to N local polytopic embeddings, whose convex-hull yields the required polytopic overapproximation on the interval $[T_1, T_2]$. The advantages of this kind of refining process, that is inspired by [64], are discussed in details below via the following claim.

Claim 5.1. Let $A \in \mathbb{R}^{n\times n}$ be a given matrix, and let $T_1 < T_2$ be given real scalars. Let $\Omega = \mathrm{co}\{X_1, X_2, \dots, X_\nu\}$, where $X_1, X_2, \dots, X_\nu$ are matrices obtained as in (5.58) on the interval $[T_1, T_2]$.
Let $\mathcal{I}_1, \mathcal{I}_2 \dots, \mathcal{I}_N$ be N compact intervals such that $\mathrm{length}(\mathcal{I}_k) = \frac{T_2 - T_1}{N}$ and $[T_1, T_2] = \bigcup_{k=1}^{N} \mathcal{I}_k$. For $k = 1, 2, \dots, N$, let

$$\Lambda_k = \mathrm{co}\{X_1^{(k)}, X_2^{(k)}, \dots, X_\nu^{(k)}\}$$

be the matrices obtained as in (5.58) on the interval $\mathcal{I}_k$. Then,

$$e^{A[T_1,T_2]} \subseteq \mathrm{co}\left\{\bigcup_{k=1}^{N} \Lambda_k\right\} \subseteq \Omega.$$

Proof. First of all notice that

$$e^{A[T_1,T_2]} = \bigcup_{k=1}^{N} e^{A\mathcal{I}_k} \subseteq \bigcup_{k=1}^{N} \Lambda_k.$$

Moreover, since $\mathcal{I}_k \subset [T_1, T_2]$ for every $k = 1, \dots, N$, by the construction of the sets Λ_k, it follows that for each $k = 1, 2, \dots, N$, $\Lambda_k \subseteq \Omega$, which in turn yields $\bigcup_{k=1}^{N} \Lambda_k \subseteq \Omega$. By isotonicity of the convex hull operator; see [12], $\mathrm{co}\left\{\bigcup_{k=1}^{N} \Lambda_k\right\} \subseteq \mathrm{co}\{\Omega\}$. Therefore, being Ω convex, the claim is proven. ■

Remark 5.13. The above result shows an underlying monotonicity of the considered refining process. Namely, by following the same arguments as in the proof of Claim 5.1, it is not difficult to show that for every $M > N$, $e^{A[T_1,T_2]} \subseteq \mathrm{co}\left\{\bigcup_{k=1}^{M} \Lambda_k\right\} \subseteq \mathrm{co}\left\{\bigcup_{k=1}^{N} \Lambda_k\right\} \subset \Omega$. Thus, the larger the value of N the tighter the overapproximation. Nonetheless, as in general the set $e^{A[T_1,T_2]}$ is not convex, the overapproximation cannot be made arbitrarily tight by selecting a value of N arbitrarily large (the pursued approach is intrinsically conservative). To somehow formalize this aspect, one can look at the asymptotic behavior of the sequence of sets $\Gamma_k = \mathrm{co}\left\{\bigcup_{i=1}^{k} \Lambda_i\right\}$ when k goes to infinity. Specifically, consider the sequence $\{\Gamma_k\}_{k=1}^{\infty}$, and observe that as argued above, for each $k \in \mathbb{N}$, one has $\Gamma_{k+1} \subseteq \Gamma_k$. Moreover, since by construction $\Gamma_1 = \Omega$, and Ω is trivially bounded, it follows that the every element of the sequence $\{\Gamma_k\}_{k=1}^{\infty}$ is compact. From these observations, by relying on the general notions of convergence for sequence of sets; see, *e.g.*, [106], one can readily show that the considered sequence converges to a convex set. Thus, since in general the set $e^{A[T_1,T_2]}$ is not convex, one should expect that the overapproximation polytope cannot be made arbitrarily tight.

5.5 Numerical Examples

Example 5.1. To illustrate the proposed polytopic embedding technique, we consider

$$A = \begin{bmatrix} 1 & 1 \\ 0 & -2 \end{bmatrix}$$

and $v \in [0, 1.5]$. As in [64], to visualize the resulting embedding, one can consider the real Jordan form of the matrix A,

$$J = U^{-1}AU = \begin{bmatrix} -2 & 0 \\ 0 & 1 \end{bmatrix}.$$

Indeed, since

$$e^{Av} = U \begin{bmatrix} e^{-2v} & 0 \\ 0 & e^{v} \end{bmatrix} U^{-1}$$

if $\{X_1, X_2, \dots, X_\nu\}$ are matrices such that $e^{Av} \in \mathrm{co}\{X_1, X_2, \dots, X_\nu\}$, then

$$e^{Jv} \in \mathrm{co}\{U^{-1}X_1U, U^{-1}X_2U, \dots, U^{-1}X_\nu U\}.$$

Figure 5.1 reports the curve $\left(e^{Jv}(1,1), e^{Jv}(2,2)\right) : [0, 1.5] \to \mathbb{R}^2$ and different polytopic overapproximations obtained by subdividing the interval $[0, 1.5]$ in several subintervals. Figure 5.2 depicts the polytopic embedding obtained with $N = 5$ and the local embedding polytopes Λ_k, for $k = 1, 2, \dots, 5$. As expected, the larger the value of N, the tighter the overapproximation.

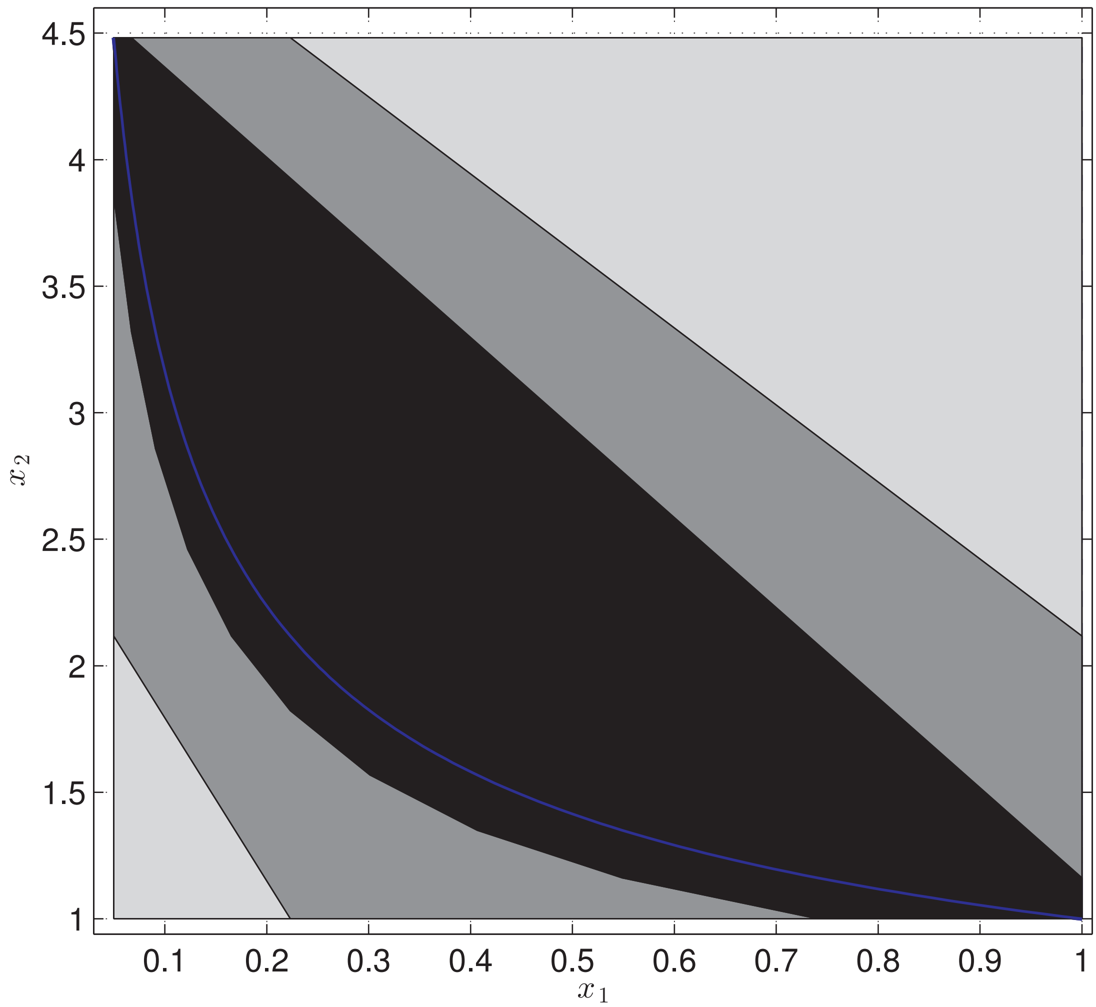


Figure 5.1: The curve $\left(e^{Jv}(1,1), e^{Jv}(2,2)\right) : [0, 1.5] \to \mathbb{R}^2$ (solid-blue) and different overapproximations, $N = 1$ (light-gray), $N = 2$ (gray), $N = 10$ (black).

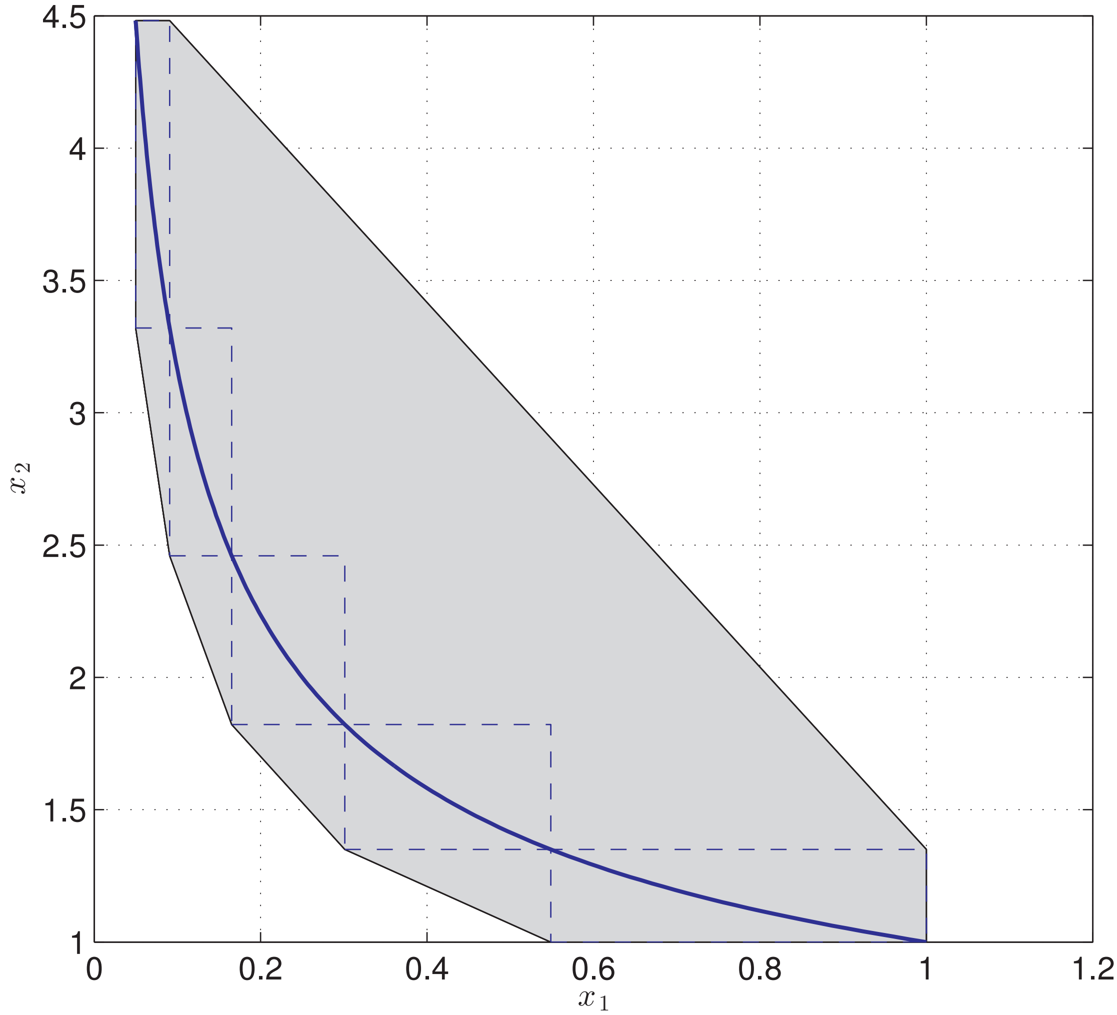


Figure 5.2: The curve $\left(e^{Jv}(1,1), e^{Jv}(2,2)\right) : [0, 1.5] \to \mathbb{R}^2$ (solid-blue), polytopic embedding with $N = 5$ (light-gray) and local polytopic overapproximations (dashed-blue).

Example 5.2. Consider the mass-spring system in [54], defined as follows

$$\dot{z} = \underbrace{\begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ -2 & 1 & -1 & 0 \\ 2 & -2 & 0 & -2 \end{bmatrix}}_{A} z + \underbrace{\begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix}}_{B} u \tag{5.59}$$

where z_1, z_2 are respectively the position of the first and the second mass, while z_3 and z_4 are respectively the speed of the first and the second mass, and u is the force applied to the second mass. Suppose that only z_1 is measurable through a biased sensor which can be accessed at most every $0.2s$ and at least every $3s$. That is, assuming the initial time $t_0 = 0$, the measured output can be expressed as

$$y(t_k) = z_1(t_k) + b \quad \forall k \in \mathbb{N}$$

where $t_1 \in [0, 3]$, $\{t_k\}_{k=1}^{\infty}$ is an increasing and unbounded sequence of positive times, such that for each $k \in \mathbb{N}$, $0.2 \le t_{k+1} - t_k \le 3$, and b is the sensor bias, *i.e.*, an unknown real constant. Notice that, the sequence $\{t_k\}_{k=1}^{\infty}$ satisfies (5.2) with $T_1 = 0.2$, and $T_2 = 3$. To fit this problem in the setting addressed by Theorem 5.1, one needs to avoid considering the bias as an external perturbation. To this end, we follow an exosystem approach, see [48, 68]. Namely, we model the constant bias affecting the output sensor as an extra state, b, such that $\dot{b} = 0$. In this way, $y = M\bar{z}$, where

$$M := \begin{bmatrix} 1 & 0 & 0 & 0 & 1 \end{bmatrix}$$

and $\bar{z} := (z, b)$. Therefore, by setting $\bar{z}$ as vector state, one can consider the extended system defined by

$$\bar{A} = \begin{bmatrix} A & 0 \\ \mathbf{0} & 0 \end{bmatrix}, \bar{B}^{\mathsf{T}} = \begin{bmatrix} B^{\mathsf{T}} & 0 \end{bmatrix}$$

that matches the class of systems considered in this paper. To build a polytopic embedding for the matrix $\bar{A}$, it suffices to build the one of A. In fact, since for each real scalar v

$$e^{\bar{A}v} = \begin{bmatrix} \mathrm{I} \\ 0 \end{bmatrix} e^{Av} \begin{bmatrix} \mathrm{I} & 0 \end{bmatrix} + \begin{bmatrix} \mathbf{0} & 0 \\ \mathbf{0} & 1 \end{bmatrix}$$

given an interval $\mathcal{I} \subset \mathbb{R}$, if for each $v \in \mathcal{I}$

$$e^{Av} \in \mathrm{co}\{X_1, X_2, \ldots, X_\nu\}$$

then by defining

$$\Theta = \begin{bmatrix} \mathrm{I} \\ \mathbf{0} \end{bmatrix}$$

it follows

$$e^{\bar{A}v} \in \left\{ \begin{bmatrix} \mathbf{0} & 0 \\ \mathbf{0} & 1 \end{bmatrix} + \mathrm{co}\{\Theta X_1 \Theta^\mathsf{T}, \Theta X_2 \Theta^\mathsf{T}, \dots, \Theta X_\nu \Theta^\mathsf{T}\} \right\} \qquad \forall v \in \mathcal{I}.$$

In particular, in this case since $\mathrm{spec}(A) = \{-0.68055 \pm 1.6332\mathrm{i}, -0.6389, -1\}$, $\nu = 16$. Once the matrices X_i are determined[3] by following the technique proposed in Section 5.4.1, via Corollary 5.1 one gets

$$P = \begin{bmatrix} 1.883 & 0.88796 & 1.3892 & 0.95109 & -1.0667 \\ 0.88796 & 12.965 & 10.415 & 10.305 & 0.091033 \\ 1.3892 & 10.415 & 10.086 & 8.6622 & -1.0351 \\ 0.95109 & 10.305 & 8.6622 & 8.8987 & -0.018634 \\ -1.0667 & 0.091033 & -1.0351 & -0.018634 & 7.6949 \end{bmatrix}$$

$$L = \begin{bmatrix} 0.77524 \\ 0.18123 \\ -0.12123 \\ -0.17406 \\ 0.22469 \end{bmatrix}.$$

Figure 5.3 reports the function $v \mapsto \lambda_{\max}\left((\mathrm{I} - LM)^\mathsf{T} e^{\bar{A}^\mathsf{T} v} P e^{\bar{A}v}(\mathrm{I} - LM) - P\right)$ as $v \in [T_1, T_2]$. As expected, the proposed design ensures that (5.12) holds.

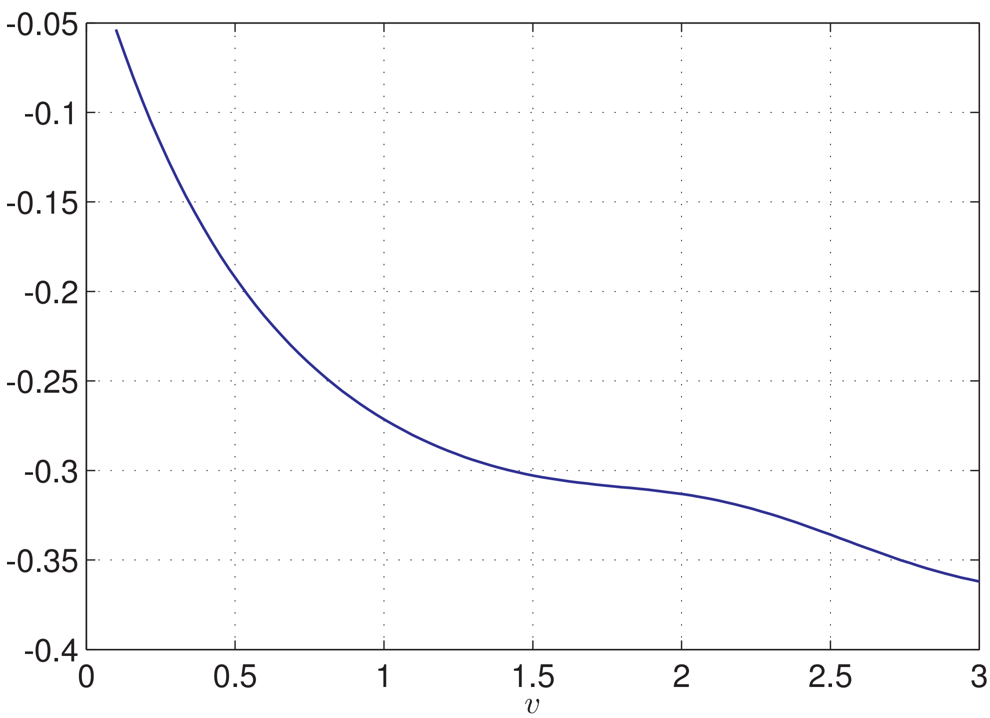


Figure 5.3: The function $v \mapsto \lambda_{\max}\left((\mathrm{I} - LM)^\mathsf{T} e^{\bar{A}^\mathsf{T} v} P e^{\bar{A}v}(\mathrm{I} - LM) - P\right)$ versus v.

[3] Such matrices are reported in Appendix B

Assume $u(t) = \sin(t)$, $b = 1$, and denote the estimate provided by the observer as $\hat{z}_e := (\hat{z}, \hat{b})$. Figure 5.4 shows the evolution of the plant state and of its estimate projected onto ordinary time. Figure 5.5 reports the evolution of the bias $\hat{b}$ projected onto ordinary time. The figures show that the designed observer reconstructs the plant state despite the presence of the sensor bias.

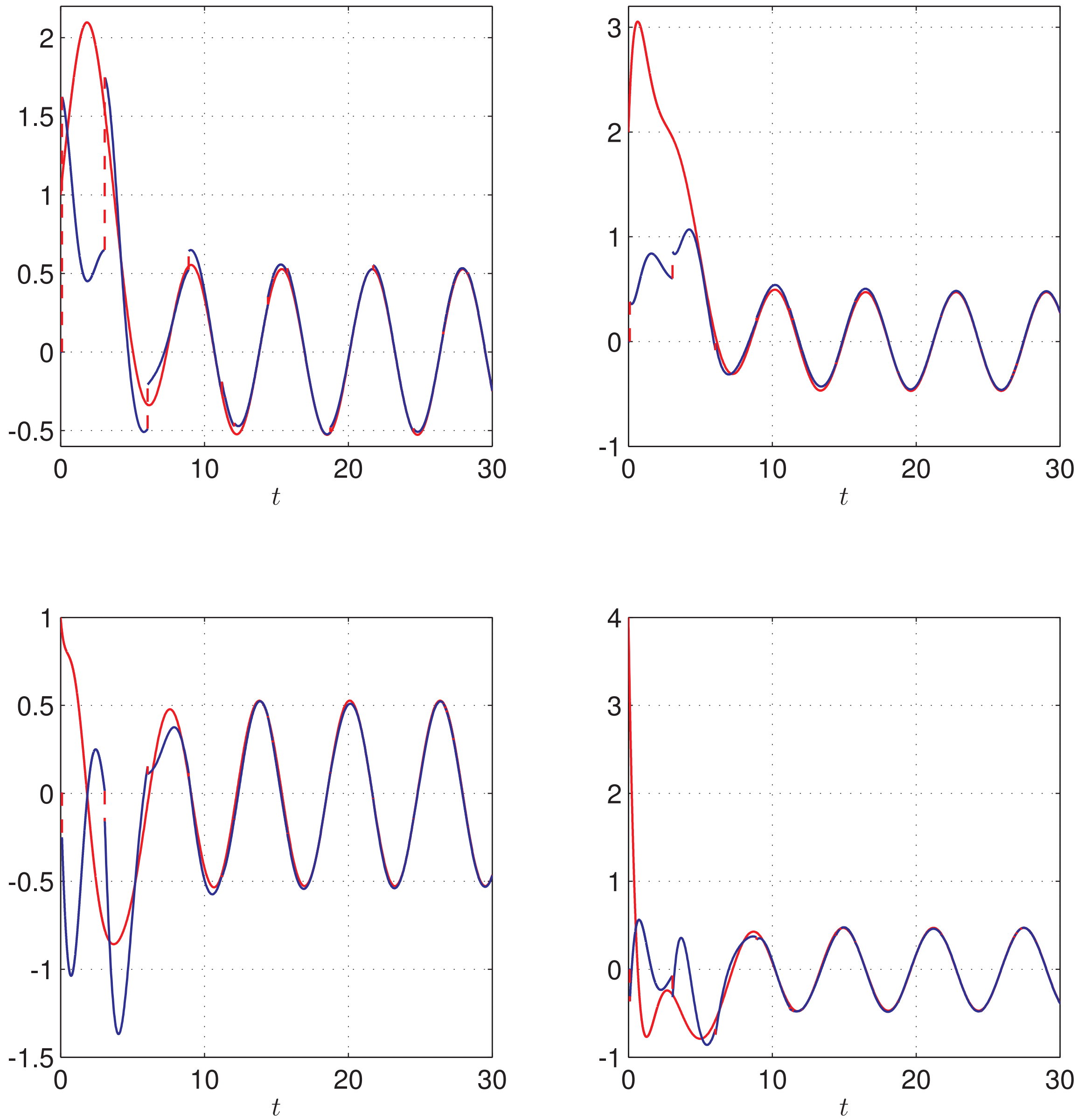


Figure 5.4: The evolution of the states z (red) and $\hat{z}$ (blue) projected onto ordinary time t.

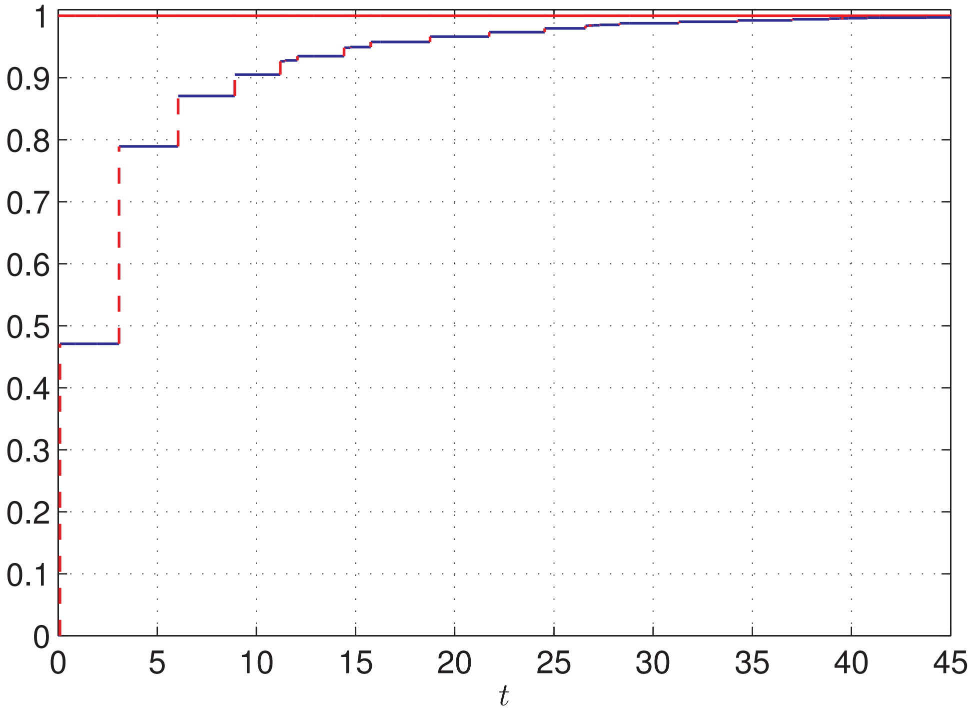


Figure 5.5: The bias b (red) and the evolution of its estimate $\hat{b}$ (blue) projected onto ordinary time t.

5.6 Comments and Conclusion

In this chapter, we proposed a methodology to model and design, through the solution to certain LMIs, a measurement-triggered observer to estimate the state of a linear plant in the presence of sporadically available measurements. The considered observer is shown to be ISS with respect to measurement noise. As shown, the effective design of the observer requires *a priori* the solution of an infinite number of LMIs, which is in practice undoable. To overcome this problem, via the introduction of a novel polytopic embedding for the exponential matrix, we embedded the obtained conditions in a polytope making the design possible via the solution to a finite number of LMIs. The proposed embedding technique somehow provides a systematic way to build a polytopic embedding for the exponential matrix, pursuing an approach analogous to the one in [29]. Hence, such a methodology is a worthwhile contribution in itself and worth of further investigations. Finally, the effectiveness of the proposed methodology is displayed in two numerical examples.

The results presented in this chapter show that the hybrid systems framework proposed in [56] permits to model and analyze the considered observer. In particular, exponential state estimation and ISS with respect to measurement noise via Lyapunov arguments were proved. Alternative frameworks, as the ones based on impulsive dynamical systems; see, *e.g.*, [103] could be used to come up with similar sufficient conditions as the ones proposed in this chapter. Another alternative approach that could be followed to tackle the problem in this chapter is the discrete-time approach considered in the literature of networked and sampled-data control systems; see [29] and the references therein. This approach consists of three stages. As a first step, a discrete-time model of the considered system is built by integration of the continuous time-dynamics in between sampling times. As a second step, asymptotic stability is established for the discretized model obtained throughout the first step. Finally, the proper intersample behavior is guaranteed by relating the continuous-time behavior with the behavior at the sampling times via the derivation of certain bounds. Following this approach, in the specific case considered in this chapter, would allow to recover some of the results presented, and also to exploit tools deriving from the literature of uncertain discrete-time systems, as, *e.g.*, polytopic Lyapunov functions ([33]), which can potentially lead to less conservative results. On the other hand, the aforementioned strategy consisting of overlooking the intersample behavior contrasts with the spirit of the hybrid system framework in [56], which studies hybrid dynamics in their whole. Then, in this setting, adopting tools from the literature of uncertain discrete-time system does not appear a viable solution. However, we would like to point out that addressing the considered problem through the hybrid system framework in [56] has several advantages. The first one is that the hybrid systems approach does not require the integration of the estimation error dynamics in between jumps. Thus, the proposed methodology can be extended to deal with more complex plants without the need of resorting on different models and frameworks. Moreover, the pursued approach, enabling the search of alternative Lyapunov functions, could be used to come up with simpler design procedures avoiding the use of the exponential matrix, which is undoable following a discrete-time approach.

The second one is that our analysis leads straight to an explicit exponential bound on the estimation error and not for a discretized version of it. Moreover, such a bound can be easily determined via the tools presented in this chapter. In this sense, our methodology allows to derive constructive results to effectively determine an exponential bound on the error trajectories in their whole. The derivation of such bounds appears intricate and far from systematic via the tools in [29].

The third one is that the hybrid system framework in [56] allows to tackle problems arising from more involved settings, where, *e.g.*, state estimation in the presence of sporadic measurements is one of the task needed to exhibit a solution to the considered problem and not the unique. This aspect will be clarified and made more concrete later in this dissertation, when such an observer will be used to build up an observer-based controller in the presence of both sporadic measurements and actuation.

Another interesting aspect that we would like to emphasize concerns with the possibility of using another modeling technique of time-triggering phenomenon presented in this chapter. Specifically, a modeling strategy similar to the one in [19] could be used to retrace the same steps illustrated within this chapter. Nevertheless, it is interesting to observe that the modeling we considered lends itself to an easy implementations in the hybrid simulator [108] than the one in [19].

Several directions of research still need to be investigated. Among them, an interesting issue concerns the construction of a measurement-triggered observer to estimate the state of more general plants, as plants characterized by sector nonlinearities. Going in that direction would allow to build interesting links with the works in [3] and the references therein. However, such an extension appears nontrivial due to the choice we considered in this chapter for the Lyapunov function, which is tailored to the linear dynamics of the plant.

Another interesting future outlook concerns the evaluation of the performances, in terms of convergence speed, offered by the proposed observer compared with observer schemes derived via emulation approach as the ones in [100]. Indeed, the main peculiarity of the scheme we considered in this chapter is that at every jump the whole state of the observer is reset. These instantaneous changes in the observer dynamics can potentially lead to an improvement of the convergence rate, while avoiding the need of a large observer gain, which is typically unwanted in practice to limit the effect of measurement noise.

Furthermore, one may envision to investigate the impact of quantized measurements on the estimation error dynamics. In particular, according to the general philosophy illustrated in [82], the ISS property shown for the estimation error dynamics with respect to measurement noise suggests that the considered observer owns the robustness needed to withstand quantized measurements.

This page is intentionally left blank.

6

A HYBRID OBSERVER WITH A CONTINUOUS INTERSAMPLE INJECTION IN THE PRESENCE OF SPORADIC MEASUREMENTS

"Experience is simply the name we give our mistakes".

– Oscar Wilde

6.1 Introduction

In this chapter, we address again the problem of exponentially estimating the state of a linear time-invariant system in the presence of sporadically available measurements. Differently from Chapter 5, we adopt an observer with a continuous-time intersample injection term. Such an intersample injection is provided by a linear dynamical system, whose state is reset to the measured output estimation error whenever a new measurement is available. The resulting system is augmented with a timer triggering the arrival of a new measurement and analyzed in a hybrid system framework. The design of the observer is performed to achieve global exponential stability of a set wherein the estimation error is equal to zero. Moreover, four computationally tractable procedures are illustrated to design the observer. Such procedures lead to four different strategies to build the proposed observer. Finally, the effectiveness of the proposed methodology is shown in two examples. Some of the results illustrated in this chapter can be found in [41].

6.2 Problem Statement

6.2.1 System Description

We consider continuous-time linear time-invariant systems of the form

$$\begin{aligned} \dot{z} &= Az \\ y &= Mz \end{aligned} \tag{6.1}$$

where $z \in \mathbb{R}^n$ and $y \in \mathbb{R}^q$ are, respectively, the state and the measured output of the system, while A and M are constant matrices of appropriate dimensions. We want to solve the same problem considered in Chapter 5 by means of an alternative observation scheme. Here below we recall the problem we solve. Assuming the initial time $t_0 = 0$, our goal is to design an observer providing an asymptotic estimate $\hat{z}$ of the state z with sporadic measurements of y. Namely, we assume that the whole output y is available only at some time instances t_k, $k \in \mathbb{N}$, not known *a priori*.

Remark 6.1. In this chapter we consider unforced plants. Whenever the considered plant is forced by an exogenous signal and such a signal is known, the results presented in this chapter apply *mutatis mutandis*. Such an assumption about the knowledge of the plant input has been already discussed in Chapter 5; see Remark 5.1.

We assume that the sequence $\{t_k\}_{k=1}^{\infty}$ is strictly increasing and unbounded, and that for such a sequence there exist two positive real scalars $T_1 \leq T_2$ such that

$$\begin{aligned} &0 \leq t_1 \leq T_2 \\ &T_1 \leq t_{k+1} - t_k \leq T_2 \qquad \forall k \in \mathbb{N}. \end{aligned} \tag{6.2}$$

As also pointed out in [64], the lower bound in condition (6.2) prevents the existence of accumulation points in the sequence $\{t_k\}_{k=1}^{\infty}$, and, hence, avoids the existence of Zeno behaviors, which are typically undesired in practice. In fact, T_1 defines a strictly positive minimum time in between consecutive measurements. Furthermore, T_2 defines the maximum sampling interval.

Since measurements of the output y are available in an impulsive fashion, assuming that the arrival of a new measurement can be instantaneously detected, to solve the considered estimation problem, inspired from [73, 100, 104], we propose the following observer with jumps

$$\left\{ \begin{array}{l} \left. \begin{array}{rcl} \dot{\hat{z}}(t) & = & A\hat{z}(t) + L\theta(t) \\ \dot{\theta}(t) & = & H\theta(t) \end{array} \right\} \ \forall t \neq t_k, k \in \mathbb{N} \\ \\ \left. \begin{array}{rcl} \hat{z}(t^+) & = & \hat{z}(t) \\ \theta(t^+) & = & y(t) - M\hat{z}(t) \end{array} \right\} \ \forall t = t_k, k \in \mathbb{N} \end{array} \right. \tag{6.3}$$

where L and H are real matrices of appropriate dimensions to be designed. The operating

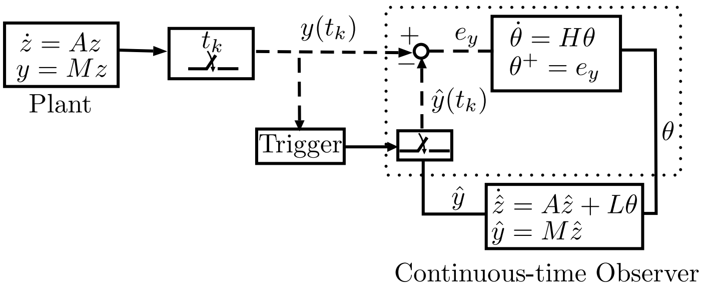


Figure 6.1: The proposed observer scheme. The dashed arrows denote impulsive data streams, while the solid arrows denote continuous data streams.

principle of the observer in (6.3) is as follows. The arrival of a new measurement triggers an instantaneous jump in the observer state. Specifically, at each jump, the measured output estimation error, *i.e.*, $y - M\hat{z}$, is instantaneously stored in θ. Then, in between consecutive measurements, θ is continuously updated according to linear continuous-time dynamics, and its value is continuously used as an intersample injection to feed a continuous-time observer; see Figure 6.1. Along the lines of [109], we formulate the state estimation problem as a set stabilization problem. Namely, our goal is to design the matrices L and H such that the set wherein the plant state z and its estimate $\hat{z}$ coincide is globally exponentially stable for the plant (6.1) interconnected with the observer in (6.3). At this stage, we define the following change of variables

$$\begin{aligned}\varepsilon &:= z - \hat{z}\\ \tilde{\theta} &:= M(z - \hat{z}) - \theta\end{aligned}$$

which defines, respectively, the estimation error and the difference between the output estimation error and θ. Hence, the two error dynamics are given by the following dynamical system with jumps:

$$\begin{cases} \begin{bmatrix}\dot{\varepsilon}(t)\\ \dot{\tilde{\theta}}(t)\end{bmatrix} = \mathcal{F}\begin{bmatrix}\varepsilon(t)\\ \tilde{\theta}(t)\end{bmatrix} & \forall t \neq t_k, k \in \mathbb{N}\\ \begin{bmatrix}\varepsilon(t^+)\\ \tilde{\theta}(t^+)\end{bmatrix} = \mathcal{G}\begin{bmatrix}\varepsilon(t)\\ \tilde{\theta}(t)\end{bmatrix} & \forall t = t_k, k \in \mathbb{N}\end{cases} \tag{6.4}$$

where

$$\begin{aligned}\mathcal{F} &:= \begin{bmatrix} A - LM & L\\ MA - MLM - HM & ML + H\end{bmatrix}\\ \mathcal{G} &:= \begin{bmatrix}\mathrm{I} & \mathbf{0}\\ \mathbf{0} & \mathbf{0}\end{bmatrix}.\end{aligned} \tag{6.5}$$

Notice that, in view of the linearity of the plant (6.1), the error dynamics are decoupled from the plant dynamics. Then, for the purpose of estimation, one can effectively only consider system (6.4).

6.2.2 Hybrid Modeling

The fact that the observer experiences jumps when a new measurement is available and evolves according to a differential equation in between updates suggests that the updating process of the error dynamics can be described via a hybrid system. Due to this, we represent the whole system composed by the plant (6.1), the observer (6.3), and the logic triggering jumps as a hybrid system. The proposed hybrid systems approach requires to model the hidden time-driven mechanism triggering the jumps of the observer. To this end, as already illustrated in Chapter 5, we augment the state of the system with an auxiliary timer variable τ that keeps track of the duration of flows and triggers a jump whenever a certain condition is verified. This additional state allows to describe the time-driven triggering mechanism as a state-driven triggering mechanism, which leads to a model that can be efficiently represented by relying on the framework for hybrid systems proposed in [56]. More precisely, we make τ to decrease as ordinary time t increases and, whenever $\tau = 0$, reset it to any point in $[T_1, T_2]$, so as to enforce (6.2). After each jump, we require the system to flow again. The whole system composed by the states ε and $\tilde{\theta}$, and the timer variable τ can be represented by the following hybrid system, which we denote $\mathcal{H}$:

$$
\mathcal{H}\begin{cases}
\left.\begin{aligned}\begin{bmatrix}\dot{\varepsilon}\\ \dot{\tilde{\theta}}\end{bmatrix} &= \mathcal{F}\begin{bmatrix}\varepsilon\\ \tilde{\theta}\end{bmatrix}\\ \dot{\tau} &= -1\end{aligned}\right\} & (\varepsilon, \tilde{\theta}, \tau) \in C\\[2ex]
\left.\begin{aligned}\begin{bmatrix}\varepsilon^+\\ \tilde{\theta}^+\end{bmatrix} &= \mathcal{G}\begin{bmatrix}\varepsilon\\ \tilde{\theta}\end{bmatrix}\\ \tau^+ &\in [T_1, T_2]\end{aligned}\right\} & (\varepsilon, \tilde{\theta}, \tau) \in D
\end{cases}
\tag{6.6a}
$$

where the flow set and the jump set are defined as

$$
\begin{aligned}
C &= \mathbb{R}^{n+q} \times [0, T_2]\\
D &= \mathbb{R}^{n+q} \times \{0\}.
\end{aligned}
\tag{6.6b}
$$

The set-valued jump map allows to capture all possible sampling events occurring within T_1 or T_2 units of time from each other. Specifically, the hybrid model in (6.6) is able to characterize not only the behavior of the analyzed system for a given sequence $\{t_k\}_{k=1}^{\infty}$, but for any sequence satisfying (6.2). We denote the state of $\mathcal{H}$ by

$$
x = (\varepsilon, \tilde{\theta}, \tau)
$$

and by f and G, respectively, the flow map and the jump map, *i.e.*,

$$
f(x) = \begin{bmatrix}\mathcal{F}\begin{bmatrix}\varepsilon\\ \tilde{\theta}\end{bmatrix}\\ -1\end{bmatrix} \qquad \forall x \in C
\tag{6.7a}
$$

$$G(x) = \begin{bmatrix} \mathcal{G}\begin{bmatrix}\varepsilon\\ \tilde{\theta}\end{bmatrix} \\ [T_1, T_2] \end{bmatrix} \qquad \forall x \in D. \tag{6.7b}$$

Then, by introducing the set[1]

$$\mathcal{A} = \{0\} \times \{0\} \times [0, T_2] \tag{6.8}$$

the problem to solve is formulated as follows:
Problem 6.1. Given the matrices A and M of appropriate dimensions and two positive scalars $T_1 \leq T_2$, design the matrices $L \in \mathbb{R}^{n\times q}$ and $H \in \mathbb{R}^{q\times q}$ such that the set $\mathcal{A}$ defined in (6.8) is GES for the hybrid system (6.6).

Concerning the existence of solutions to system (6.6), by relying on the concept of solution proposed in Definition 4.5, it is straightforward to check that for every initial condition $\phi(0,0) \in C \cup D$ every maximal solution to (6.6) is complete ensuring that the estimation error approaches zero when $t+j$ goes to infinity. Thus, completeness of the maximal solutions to (6.6), as required in the statement of Problem 6.1, is guaranteed for any choice of the gains L and H. In addition, we can characterize the domain of these solutions. Indeed, as in Chapter 5, for every initial condition $\phi(0,0) \in C \cup D$, the domain of every maximal solution ϕ to (6.6) can be written as follows:

$$\operatorname{dom}\phi = \bigcup_{j\in\mathbb{N}_0} ([t_j, t_{j+1}]) \times \{j\} \tag{6.9}$$

with $t_0 = 0$ and

$$\begin{aligned} &T_1 \leq t_{j+1} - t_j \leq T_2 \qquad \forall j \in \mathbb{N} \\ &0 \leq t_1 \leq T_2 \end{aligned} \tag{6.10}$$

where $\operatorname{dom}\phi$ is the domain of the solution ϕ, which is a hybrid time domain. Therefore, the structure of the above hybrid time domain implies that for each $(t, j) \in \operatorname{dom}\phi$ we have

$$t \geq T_1 j - T_1 \tag{6.11}$$

the latter relation will play a fundamental role in establishing GES of $\mathcal{A}$ for hybrid system (6.6).

6.3 Preliminary Results

6.3.1 Conditions for GES

In this section we provide a first sufficient condition to solve Problem 6.1. Such a condition is obtained by the adoption of a Lyapunov-like function, that is inspired by [47, 55]. To

[1] By the definition of system (6.6) and of the set $\mathcal{A}$, for every $x \in C \cup D \cup G(D)$, $|x|_{\mathcal{A}} = \|(\varepsilon, \tilde{\theta})\|$.

pursue this approach, let us consider the following assumption, whose role will be clarified right after via Theorem 6.1.

Assumption 6.1. Consider (6.7a) and set

$$\mathcal{F} = \begin{bmatrix} \mathcal{F}_{11} & \mathcal{F}_{12} \\ \mathcal{F}_{21} & \mathcal{F}_{22} \end{bmatrix}.$$

There exist two continuously differentiable functions $V_1 \colon \mathbb{R}^n \to \mathbb{R}$, $V_2 \colon \mathbb{R}^q \to \mathbb{R}$, positive real scalars $\alpha_1, \alpha_2, \omega_1, \omega_2, \sigma, \lambda_c$ such that for each $(\varepsilon, \tilde{\theta}, \tau) \in C$

(A1) $\alpha_1 \|\varepsilon\|^2 \leq V_1(\varepsilon) \leq \alpha_2 \|\varepsilon\|^2$

(A2) $\omega_1 \|\tilde{\theta}\|^2 \leq V_2(\tilde{\theta}) \leq \omega_2 \|\tilde{\theta}\|^2$

(A3) $\langle \nabla V_1(\varepsilon), \mathcal{F}_{11}\varepsilon + \mathcal{F}_{12}\tilde{\theta}\rangle + e^{\sigma\tau}\langle \nabla V_2(\tilde{\theta}), \mathcal{F}_{22}\tilde{\theta} + \mathcal{F}_{21}\varepsilon\rangle - \sigma e^{\sigma\tau} V_2(\tilde{\theta}) \leq -\lambda_c(\|\varepsilon\|^2 + \|\tilde{\theta}\|^2)$

$\triangle$

Sufficient conditions to let Assumption 6.1 hold will be given in the sequel of this chapter.

Theorem 6.1. Let Assumption 6.1 hold. Then the set $\mathcal{A}$ defined in (6.8) is GES for hybrid system (6.6).

The proof of the above theorem requires the following lemma, whose proof is given later.

Lemma 6.1. Let λ_c be any strictly positive real number. Pick

$$\lambda \in \left(0, \frac{\lambda_c T_1}{1 + T_1}\right], \omega \geq \lambda.$$

Let ϕ be a solution to the hybrid system (6.6). Then, for every $(t, j) \in \operatorname{dom} \phi$, one has

$$-\lambda_c t \leq \omega - \lambda(t + j). \tag{6.12}$$

Now we are in position to prove Theorem 6.1

Proof of Theorem 6.1. Inspired by [55, Example 27], consider the following Lyapunov function candidate for the hybrid system (6.6) defined for every $x \in \mathbb{R}^{n+q} \times \mathbb{R}_{\geq 0}$:

$$V(x) = V_1(\varepsilon) + e^{\sigma\tau} V_2(\tilde{\theta}). \tag{6.13}$$

To prove the claim, we rely on the proof of the stability result provided in [56, Proposition 3.29]. To this end, notice that by setting $\rho_1 = \min\{\alpha_1, \omega_1\}$ and $\rho_2 = \max\{\alpha_2, \omega_2 e^{\sigma T_2}\}$, in view of the definition of the set $\mathcal{A}$ in (6.8), one gets

$$\rho_1 |x|^2_{\mathcal{A}} \leq V(x) \leq \rho_2 |x|^2_{\mathcal{A}} \qquad \forall x \in C \cup D \cup G(D). \tag{6.14}$$

By straightforward calculations, and from the definition of the flow map f in (6.7a), for each

$x \in C$, one has

$$\langle \nabla V(x), f(x) \rangle = \langle \nabla V_1(\varepsilon), \mathcal{F}_{11}\varepsilon + \mathcal{F}_{12}\tilde{\theta} \rangle + e^{\sigma\tau} \langle \nabla V_2(\tilde{\theta}), \mathcal{F}_{22}\tilde{\theta} + \mathcal{F}_{21}\varepsilon \rangle - \sigma e^{\sigma\tau} V_2(\tilde{\theta}).$$

Thus from Assumption 6.1, the above relation yields

$$\langle \nabla V(x), f(x) \rangle \leq -\lambda_c(\|\varepsilon\|^2 + \|\tilde{\theta}\|^2) = -\lambda_c |x|_{\mathcal{A}}^2 \qquad \forall x \in C \tag{6.15}$$

which in turn thanks to (6.14) gives

$$\langle \nabla V(x), f(x) \rangle \leq -\frac{\lambda_c}{\rho_2} V(x) \qquad \forall x \in C. \tag{6.16}$$

Now, notice that for every $g \in G(x)$, there exists a real scalar v belonging to the interval $[T_1, T_2]$ such that $g = (\varepsilon, \mathbf{0}, v)$. Then, for every $g \in G(x)$ and for every $x \in D$, one has

$$V(g) - V(x) = -V_2(\tilde{\theta}) \leq 0. \tag{6.17}$$

Pick

$$\omega = \lambda = \frac{\lambda_c T_1}{\rho_2 \alpha_2 (1 + T_1)}$$

and let ϕ be a maximal solution to (6.6). As shown in the proof of [56, Proposition 3.29], thanks to (6.16) and (6.17), direct integration of $(t, j) \mapsto V(\phi(t, j))$ over $\operatorname{dom} \phi$ yields

$$V(\phi(t, j)) \leq e^{-\frac{\lambda_c}{\rho_2} t} V(\phi(0, 0)). \tag{6.18}$$

Then, due to the choice operated for λ according to Lemma 6.1, from (6.18), it follows that

$$V(\phi(t, j)) \leq e^{-\lambda(t+j)} e^{\lambda} V(\phi(0, 0)) \qquad \forall (t, j) \in \operatorname{dom} \phi. \tag{6.19}$$

Still, in view of (6.14), one has

$$|\phi(t, j)|_{\mathcal{A}} \leq e^{-\frac{\lambda(t+j)}{2}} e^{\frac{\omega}{2}} \left(\frac{\rho_2}{\rho_1} \right)^{\frac{1}{2}} |\phi(0, 0)|_{\mathcal{A}} \qquad \forall (t, j) \in \operatorname{dom} \phi \tag{6.20}$$

which implies that the set $\mathcal{A}$ defined in (6.8) is GES for system (6.6). ■

Now, the proof of Lemma 6.1 is given

Proof of Lemma 6.1. From (6.12), by rearranging the terms, one gets

$$(-\lambda_t + \lambda)t + \lambda j - \omega \leq 0. \tag{6.21}$$

Now, pick any solution ϕ to hybrid system (6.6). From (6.11), it follows that for every $(t, j) \in \operatorname{dom} \phi$

$$j \leq \frac{t}{T_1} + 1 \tag{6.22}$$

then, for every strictly positive scalar λ, from the latter expression and for every $(t,j) \in \operatorname{dom}\phi$, one gets

$$(-\lambda_t + \lambda)t + \lambda j - \omega \leq \left(-\lambda_t + \lambda + \frac{\lambda}{T_1}\right) t + \lambda - \omega. \tag{6.23}$$

Thus, being T_1 strictly positive, by selecting

$$\lambda \in \left(0, \frac{\lambda_t T_1}{1+T_1}\right], \omega \geq \lambda$$

yields (6.21), which concludes the proof. ∎

Theorem 6.1 shows that if there exist matrices $L \in \mathbb{R}^{n\times q}$ and $H \in \mathbb{R}^{q\times q}$ such that Assumption 6.1 holds, then such matrices are a solution to Problem 6.1. Next, we provide two alternative sufficient conditions ensuring the satisfaction of Assumption 6.1.

Proposition 6.1. *Consider* (6.7a) *and set*

$$\mathcal{F} = \begin{bmatrix} \mathcal{F}_{11} & \mathcal{F}_{12} \\ \mathcal{F}_{21} & \mathcal{F}_{22} \end{bmatrix}.$$

If there exist two continuously differentiable functions $V_1 \colon \mathbb{R}^n \to \mathbb{R}$ and $V_2 \colon \mathbb{R}^q \to \mathbb{R}$, positive real scalars $\alpha_1, \alpha_2, \beta, \gamma, \omega_1, \omega_2, \rho, \xi, \sigma$ such that

(i) $\alpha_1\|\varepsilon\|^2 \leq V_1(\varepsilon) \leq \alpha_2\|\varepsilon\|^2 \quad \forall \varepsilon \in \mathbb{R}^n$

(ii) $\langle \nabla V_1(\varepsilon), \mathcal{F}_{11}\varepsilon + \mathcal{F}_{12}\tilde{\theta}\rangle \leq -\beta\|\varepsilon\|^2 + \gamma\|\tilde{\theta}\|^2 \quad \forall (\varepsilon, \tilde{\theta}) \in \mathbb{R}^n \times \mathbb{R}^q$

(iii) $\omega_1\|\tilde{\theta}\|^2 \leq V_2(\tilde{\theta}) \leq \omega_2\|\tilde{\theta}\|^2 \quad \forall \tilde{\theta} \in \mathbb{R}^q$

(iv) $\langle \nabla V_2(\tilde{\theta}), \mathcal{F}_{22}\tilde{\theta} + \mathcal{F}_{21}\varepsilon\rangle \leq \rho\|\varepsilon\|^2 + \xi\|\tilde{\theta}\|^2 \quad \forall (\varepsilon, \tilde{\theta}) \in \mathbb{R}^n \times \mathbb{R}^q$

such that

$$\sigma\omega_1 - \gamma > 0 \tag{6.24a}$$

$$T_2 < \frac{1}{\sigma} \ln\left(\min\left\{\frac{\beta}{\rho}, \frac{\sigma\omega_1 - \gamma}{\xi}\right\}\right). \tag{6.24b}$$

Then Assumption 6.1 holds, with

$$\lambda_c = \min\left\{|-\beta + e^{\sigma T_2}\rho|, |\gamma + e^{\sigma T_2}\xi - \sigma\omega_1|\right\}. \tag{6.25}$$

Proof. From (ii) and (iv), it follows that for each $(\varepsilon, \tilde{\theta}, \tau) \in \mathbb{R}^n \times \mathbb{R}^q \times [0, T_2]$

$$\begin{aligned} &\langle \nabla V_1(\varepsilon), \mathcal{F}_{11}\varepsilon + \mathcal{F}_{12}\tilde{\theta}\rangle + e^{\sigma\tau}\langle \nabla V_2(\tilde{\theta}), \mathcal{F}_{22}\tilde{\theta} + \mathcal{F}_{21}\varepsilon\rangle - \sigma e^{\sigma\tau} V_2(\tilde{\theta}) \leq \\ &-\beta\|\varepsilon\|^2 + \gamma\|\tilde{\theta}\|^2 + e^{\sigma\tau}(\rho\|\varepsilon\|^2 + \xi\|\tilde{\theta}\|^2) - \sigma e^{\sigma\tau} V_2(\tilde{\theta}). \end{aligned} \tag{6.26}$$

By using (iii) and by rearranging the terms, from the above inequality one gets

$$\begin{aligned}&\langle \nabla V_1(\varepsilon), \mathcal{F}_{11}\varepsilon + \mathcal{F}_{12}\tilde{\theta}\rangle + e^{\sigma\tau}\langle \nabla V_2(\tilde{\theta}), \mathcal{F}_{22}\tilde{\theta} + \mathcal{F}_{21}\varepsilon\rangle - \sigma e^{\sigma\tau} V_2(\tilde{\theta}) \leq \\ &\underbrace{(-\beta + e^{\sigma T_2}\rho)}_{\lambda_1}\|\varepsilon\|^2 + \underbrace{(\gamma + e^{\sigma T_2}\xi - \sigma\omega_1)}_{\lambda_2}\|\tilde{\theta}\|^2.\end{aligned} \tag{6.27}$$

Notice that λ_1 and λ_2 are strictly negative due to (6.24b). Thus, by setting $\lambda_c = \min\{|\lambda_1|, |\lambda_2|\}$, the above result is proven. ■

Conditions (i)-(iv) in Proposition 6.1 are rather mild to satisfy. In particular, by selecting L such that $A - LM$ is Hurwitz, (i)-(ii) can be always satisfied by selecting $V_1(\varepsilon) = \varepsilon^{\mathsf{T}} P_1 \varepsilon$, with $P_1 \in \mathcal{S}_+^n$ and such that $\mathrm{He}(P_1(A - LM)) < \mathbf{0}$. (iii)-(iv) can be always satisfied, *e.g.*, by selecting for V_2 any positive definite quadratic function. The most challenging issue consists of fulfilling (6.24). In particular, due to $T_2 > 0$ a necessary condition for the applicability of Proposition 6.1 is that

$$\begin{aligned}&\beta > \rho\\ &\sigma\omega_1 - \gamma - \xi > 0.\end{aligned}$$

Moreover, given positive scalars $\alpha_1, \alpha_2, \beta, \gamma, \omega_1, \omega_2$ satisfying (i), (ii), (iii), (iv), the satisfaction of (6.24a) can be ensured by selecting σ large enough. However, notice that

$$\lim_{\sigma\to\infty} \frac{1}{\sigma} \ln\left(\min\left\{\frac{\beta}{\rho}, \frac{\sigma\omega_1 - \gamma}{\xi}\right\}\right) = 0$$

therefore, enlarging the value of σ may prevent from fulfilling (6.24b).

To somehow overcome this problem, as follows we provide an alternative sufficient condition to let Assumption 6.1 hold.

Proposition 6.2. *Consider* (6.7a) *and set*

$$\mathcal{F} = \begin{bmatrix} \mathcal{F}_{11} & \mathcal{F}_{12} \\ \mathcal{F}_{21} & \mathcal{F}_{22} \end{bmatrix}.$$

If there exist two continuously differentiable functions $V_1\colon \mathbb{R}^n \to \mathbb{R}$, and $V_2\colon \mathbb{R}^q \to \mathbb{R}$, positive real scalars $\alpha_1, \alpha_2, \beta, \gamma, \omega_1, \omega_2, \rho, \mu, \sigma$ such that

(a) $\alpha_1\|\varepsilon\|^2 \leq V_1(\varepsilon) \leq \alpha_2\|\varepsilon\|^2 \quad \forall \varepsilon \in \mathbb{R}^n$

(b) $\langle \nabla V_1(\varepsilon), \mathcal{F}_{11}\varepsilon + \mathcal{F}_{12}\tilde{\theta}\rangle \leq -\beta\|\varepsilon\|^2 + \gamma\|\tilde{\theta}\|^2 \quad \forall(\varepsilon, \tilde{\theta}) \in \mathbb{R}^n \times \mathbb{R}^q$

(c) $\omega_1\|\tilde{\theta}\|^2 \leq V_2(\tilde{\theta}) \leq \omega_2\|\tilde{\theta}\|^2 \quad \forall \tilde{\theta} \in \mathbb{R}^q$

(d) $\langle \nabla V_2(\tilde{\theta}), \mathcal{F}_{22}\tilde{\theta} + \mathcal{F}_{21}\varepsilon\rangle \leq -\mu\|\tilde{\theta}\|^2 + \rho\|\varepsilon\|^2 \quad \forall(\varepsilon, \tilde{\theta}) \in \mathbb{R}^n \times \mathbb{R}^q$

and

$$\sigma\omega_1 - \gamma + \mu > 0 \tag{6.28a}$$

$$T_2 < \frac{1}{\sigma}\ln\left(\frac{\beta}{\rho}\right). \tag{6.28b}$$

Then Assumption 6.1 holds, with

$$\lambda_c = \min\left\{|-\beta + e^{\sigma T_2}\rho|, |\gamma - \mu - \sigma\omega_1|\right\}. \tag{6.29}$$

Proof. From (b) and (d), it follows that for each $(\varepsilon, \tilde{\theta}, \tau) \in \mathbb{R}^n \times \mathbb{R}^q \times [0, T_2]$

$$\begin{aligned} &\langle \nabla V_1(\varepsilon), \mathcal{F}_{11}\varepsilon + \mathcal{F}_{12}\tilde{\theta}\rangle + e^{\sigma\tau}\langle \nabla V_2(\tilde{\theta}), \mathcal{F}_{22}\tilde{\theta} + \mathcal{F}_{21}\varepsilon\rangle - \sigma e^{\sigma\tau}V_2(\tilde{\theta}) \leq \\ &-\beta\|\varepsilon\|^2 + \gamma\|\tilde{\theta}\|^2 + e^{\sigma\tau}(\rho\|\varepsilon\|^2 - \mu\|\tilde{\theta}\|^2) - \sigma e^{\sigma\tau}V_2(\tilde{\theta}). \end{aligned} \tag{6.30}$$

By using (c) and by rearranging the terms, from the above inequality one gets

$$\begin{aligned} &\langle \nabla V_1(\varepsilon), \mathcal{F}_{11}\varepsilon + \mathcal{F}_{12}\tilde{\theta}\rangle + e^{\sigma\tau}\langle \nabla V_2(\tilde{\theta}), \mathcal{F}_{22}\tilde{\theta} + \mathcal{F}_{21}\varepsilon\rangle - \sigma e^{\sigma\tau}V_2(\tilde{\theta}) \leq \\ &\underbrace{(-\beta + e^{\sigma T_2}\rho)}_{\lambda_1}\|\varepsilon\|^2 + \underbrace{(\gamma - \mu - \sigma\omega_1)}_{\lambda_2}\|\tilde{\theta}\|^2. \end{aligned} \tag{6.31}$$

Notice that λ_1 and λ_2 are strictly negative due to (6.28). Thus, by setting $\lambda_c = \min\{|\lambda_1|, |\lambda_2|\}$, the above result is proven. ■

Also in this case, due to $T_2 > 0$ a necessary condition for the applicability of Proposition 6.2 is that $\beta > \rho$. Nonetheless, differently from Proposition 6.1, due to $\mu > 0$, (6.28a) appears less stringent than (6.24a). In particular, (6.28a) can be *a priori* satisfied by selecting a smaller value for σ with respect to (6.24a). Such a benefit arises from having required in Proposition 6.2 a stronger assumption than in Proposition 6.1, namely (d). Nevertheless, such an assumption can be always satisfied. Indeed, due to the linearity of the $\tilde{\theta}$ flow dynamics, (d) turns out to be equivalent to the Hurwitzness of the matrix $\mathcal{F}_{22}$, property that can be always ensured via a suitable choice for H. Even more, due to the expression of $\mathcal{F}_{22}$, μ in (d) can be selected arbitrarily large via the selection of the matrix H. However, it is worthwhile to observe that, in general, picking for μ an overly large value may lead to a large value of ρ in (d), which in turn may render (6.28b) unfulfilled.

Proposition 6.1 and Proposition 6.2 provide first indications on how a solution to Problem 6.1 could be determined and also on the main challenges in determining such a solution. The approach presented, though leading to different conclusion, is similar to some extent to the one considered in [73, 100]. However, the use of Proposition 6.1 and Proposition 6.2 to solve Problem 6.1 entails several drawbacks. The first one concerns the fact that the results given in Proposition 6.1 and Proposition 6.2 dramatically depend on the choice performed for the two functions V_1 and V_2, and on the way the bounds given in (a)-(b)-(c)-(d) in Proposition 6.1 or in Proposition 6.2 are obtained. The second one is that Proposition 6.1 and Proposition 6.2 do not provide a clear strategy to select the two gains L and H so as to solve the considered problem for given data (A, M, T_2). In particular, as it is in [73], the proposed approach is rather cumbersome whenever one attempts to design the considered observer. Roughly speaking, Proposition 6.1 and Proposition 6.2 are essentially analysis results. Therefore, to build up an effective design strategy further work is needed.

Specifically, to overcome all the drawbacks illustrated above, we pursue a constructive approach. In particular, by restricting the search of the two functions V_1 and V_2 in Assumption 6.1 to the class of quadratic functions, as follows we provide a sufficient condition to let Assumption 6.1 hold that is based on the solution to certain matrix inequalities. Via this step, essentially we reduce the solution to Problem 6.1 to the solution to a feasibility problem of certain matrix inequalities. The solution of such a problem provides in one shot the solution to Problem 6.1. As argued in the above discussions, by selecting the two functions V_1 and V_2 in Assumption 6.1 as quadratic functions allows to fulfill (a)-(b)-(c)-(d) either in Proposition 6.1 or in Proposition 6.2. Hence although conservative, this choice appears promising to solve the considered problem.

6.4 Observer Design via Matrix Inequalities

The following result is one of the key results within this section. It turns the solution to Problem 6.1 into the solution to the feasibility problem to certain matrix inequalities.

Theorem 6.2. If there exist $P_1 \in \mathcal{S}_+^n, P_2 \in \mathcal{S}_+^q$, a positive real scalar σ, and two matrices $L \in \mathbb{R}^{n\times q}$, and $H \in \mathbb{R}^{q\times q}$ such that

$$\begin{aligned}
\mathcal{M}_1 &= \begin{bmatrix} \mathrm{He}(P_1(A-LM)) & P_1L + (MA - MLM - HM)^\mathsf{T} P_2 \\ \bullet & \mathrm{He}(P_2(ML+H)) - \sigma P_2 \end{bmatrix} < \mathbf{0} \\
\mathcal{M}_2 &= \begin{bmatrix} \mathrm{He}(P_1(A-LM)) & P_1L + e^{\sigma T_2}(MA - MLM - HM)^\mathsf{T} P_2 \\ \bullet & e^{\sigma T_2}\left(\mathrm{He}(P_2(ML+H)) - \sigma P_2\right) \end{bmatrix} < \mathbf{0}
\end{aligned} \tag{6.32}$$

then Assumption 6.1 holds.

Proof. Define for each $(\varepsilon, \tilde{\theta}) \in \mathbb{R}^n \times \mathbb{R}^q$

$$\begin{aligned}
V_1(\varepsilon) &= \varepsilon^\mathsf{T} P_1 \varepsilon \\
V_2(\tilde{\theta}) &= \tilde{\theta}^\mathsf{T} P_2 \tilde{\theta}.
\end{aligned} \tag{6.33}$$

Set

$$\begin{aligned}
\alpha_1 &= \lambda_{\min}(P_1) \\
\alpha_2 &= \lambda_{\max}(P_1) \\
\omega_1 &= \lambda_{\min}(P_2) \\
\omega_2 &= e^{\sigma T_2}\lambda_{\max}(P_2).
\end{aligned}$$

By straightforward calculations, it follows that for each $(\varepsilon, \tilde{\theta}, \tau) \in C$, and each positive real scalar σ,

$$\begin{aligned}
&\langle \nabla V_1(\varepsilon), \mathcal{F}_{11}\varepsilon + \mathcal{F}_{12}\tilde{\theta}\rangle + \langle \nabla V_2(\tilde{\theta}), \mathcal{F}_{22} + \mathcal{F}_{21}\varepsilon\rangle - \sigma e^{\sigma\tau} V_2(\tilde{\theta}) = \\
&\varepsilon^\mathsf{T}\, \mathrm{He}(P_1(A-LM))\varepsilon + 2\varepsilon^\mathsf{T} P_1 L\tilde{\theta} + e^{\sigma\tau}\tilde{\theta}^\mathsf{T}\, \mathrm{He}(P_2(ML+H))\tilde{\theta} \\
&+ 2e^{\sigma\tau}\tilde{\theta}^\mathsf{T} P_2(MA - MLM - HM)\varepsilon - \sigma e^{\sigma\tau}\tilde{\theta}^\mathsf{T} P_2\tilde{\theta}.
\end{aligned} \tag{6.34}$$

By defining the vector $\zeta = (\varepsilon, \tilde{\theta})$, the above expression can be equivalently rewritten as follows

$$\zeta^{\mathsf{T}} \underbrace{\begin{bmatrix} \mathrm{He}(P_1(A-LM)) & P_1 L + e^{\sigma\tau}(MA - MLM - HM)^{\mathsf{T}} P_2 \\ \bullet & e^{\sigma\tau}(\mathrm{He}(P_2(ML+H)) - \sigma P_2) \end{bmatrix}}_{\mathcal{M}(\tau)} \zeta. \tag{6.35}$$

Now, notice that, for any positive σ, there exists a scalar function $\xi_\sigma \colon [0, T_2] \to [0,1]$, such that for every $\tau \in [0, T_2]$, $e^{\sigma\tau} = \xi_\sigma(\tau) + (1 - \xi_\sigma(\tau)) e^{\sigma T_2}$. Thus, for each $x \in C$, (6.35) can be rewritten as

$$\zeta^{\mathsf{T}} \Big(\xi_\sigma(\tau) \mathcal{M}_1 + (1 - \xi_\sigma(\tau)) \mathcal{M}_2 \Big) \zeta \tag{6.36}$$

where $\mathcal{M}_1$ and $\mathcal{M}_2$ are defined in (6.32). Thus, in view of (6.32), for each $\tau \in [0, T_2]$,

$$\mathcal{M}(\tau) < \mathbf{0}. \tag{6.37}$$

Moreover, since $\mathcal{M}(\tau)$ depends continuously on τ, and τ belongs to a compact interval, the following bound holds

$$\mathcal{M}(\tau) \leq \max_{\tau \in [0, T_2]} \lambda_{\max}(\mathcal{M}(\tau)) \mathrm{I} \qquad \forall \tau \in [0, T_2].$$

Thus, by selecting

$$\lambda_c = - \max_{\tau \in [0, T_2]} \lambda_{\max}(\mathcal{M}(\tau))$$

which is positive due to (6.37), for each $x \in C$, from (6.34) one has

$$\langle \nabla V_1(\varepsilon), \mathcal{F}_{11}\varepsilon + \mathcal{F}_{12}\tilde{\theta} \rangle + \langle \nabla V_2(\tilde{\theta}), \mathcal{F}_{22} + \mathcal{F}_{21}\varepsilon \rangle - \sigma e^{\sigma\tau} V_2(\tilde{\theta}) \leq -\lambda_c (\|\varepsilon\|^2 + \|\tilde{\theta}\|^2). \tag{6.38}$$

Hence Assumption 6.1 holds, concluding the proof. ■

Remark 6.2. The feasibility of the conditions given in Theorem 6.2 requires a detectable pair (A, M), (though this condition is in general only necessary). It is worthwhile to remark that, differently from the observer considered in Chapter 5, *a priori*, we do not require the detectability of the pair (e^{Av}, Me^{Av}) for each v belonging to $[T_1, T_2]$, which would be a more restrictive condition.

6.5 Numerical Issues in the Solution to Problem 6.1

In the previous section a condition to guarantee GES of the set $\mathcal{A}$ for system and based on the feasibility of some matrix inequalities was provided. However, due to its form, such a condition is not computationally tractable to obtain a solution to Problem 6.1. Indeed, condition (6.32) is nonlinear in the design variables P_1, P_2, σ, H and L, so further work is needed to derive a numerically tractable design procedure for the proposed observer. Specifically, the nonlinearities present in (6.32) are due to both the bilinear terms involving the matrices P_1, P_2, L, H, and the scalar σ, as well as the fact that σ also appears in a nonlinear fashion via the exponential function. Nevertheless, from a numerical standpoint, the nonlinearities involving the scalar σ are easily manageable. Indeed, σ can be treated as a tuning parameter or being selected via a grid search. Thus, the main issue to tackle pertains to the other nonlinearities present in (6.32). To this aim, in the sequel we provide four constructive sufficient conditions to solve Problem 6.1 via the solution of the feasibility problem to certain linear matrix inequalities.

6.5.1 Two First Design Results

Proposition 6.3. *If there exist $P_1 \in \mathcal{S}_+^n, P_2 \in \mathcal{S}_+^q$, a positive real scalar σ, $J \in \mathbb{R}^{n\times q}$, and $Y \in \mathbb{R}^{q\times q}$ such that*

$$\begin{aligned}&\begin{bmatrix}\mathrm{He}(P_1A - JM) & J + A^\top M^\top P_2 - M^\top Y\\ \bullet & \mathrm{He}(Y) - \sigma P_2\end{bmatrix} < \mathbf{0}\\ &\begin{bmatrix}\mathrm{He}(P_1A - JM) & J + e^{\sigma T_2}(A^\top M^\top P_2 - M^\top Y)\\ \bullet & (\mathrm{He}(Y) - \sigma P_2)e^{\sigma T_2}\end{bmatrix} < \mathbf{0}\end{aligned} \tag{6.39}$$

then $L = P_1^{-1}J, H = P_2^{-1}Y^\top - ML$ is a solution to Problem 6.1.

Proof. By setting $H = P_2^{-1}Y^\top - ML$ and $J = P_1L$ in (6.32) yields (6.39), thus by virtue of Theorem 6.2, this concludes the proof. ∎

The main idea behind the above result consists of selecting the design variable H so as to cancel out the term MLM, which would unlikely lead to tractable conditions. Obviously other approaches can be pursued to cope with this issue.

Building on the previous result, another strategy to design the proposed observer is given next. Such a strategy leads to the well known observer scheme in [73].

Corollary 6.1. If there exist $P_1 \in \mathcal{S}_+^n, P_2 \in \mathcal{S}_+^q$, a positive real scalar σ, and $J \in \mathbb{R}^{n\times q}$ such that

$$\begin{aligned}&\begin{bmatrix}\mathrm{He}(P_1A - JM) & J + A^\top M^\top P_2\\ \bullet & -\sigma P_2\end{bmatrix} < \mathbf{0}\\ &\begin{bmatrix}\mathrm{He}(P_1A - JM) & J + e^{\sigma T_2}A^\top M^\top P_2\\ \bullet & -e^{\sigma T_2}\sigma P_2\end{bmatrix} < \mathbf{0}\end{aligned} \tag{6.40}$$

then $L = P_1^{-1}J, H = -ML$ is a solution to Problem 6.1.

Proof. The proof follows directly from Proposition 6.3 by selecting $Y = \mathbf{0}$. ∎

As mentioned above, the proposed choice for the gain H leads to the predictor-based observer scheme proposed in [72, 73], though written in different coordinates. Indeed, whenever $H = -ML$, by rewriting (6.3) via the following invertible change of variables $(\hat{z}, w) = (\hat{z}, \theta + M\hat{z})$, yields the same observer in [72, 73].

In the next sections, we present two other design procedures. The derivation of such procedures is based on an equivalent condition to Theorem 6.2, which is formulated introducing some slack variables via the use of the projection lemma; see [99].

6.5.2 Slack Variables-based Design

Before stating the main result, let us consider the following fact.
Fact 6.1. The matrix $\mathcal{F}$ in (6.5) can be factorized as follows

$$\mathcal{F} = \underbrace{\begin{bmatrix} \mathrm{I} & \mathbf{0} \\ M & \mathrm{I} \end{bmatrix}}_{\mathcal{F}_l} \underbrace{\begin{bmatrix} A - LM & L \\ -HM & H \end{bmatrix}}_{\mathcal{F}_r} \tag{6.41}$$

where $\mathcal{F}_l$ is nonsingular.

Building on this fact, the following result provides an equivalent condition to condition (6.32) in Theorem 6.2, in which the term MLM no longer appears.
Corollary 6.2. Let $P_1 \in \mathcal{S}^n_+$, $P_2 \in \mathcal{S}^q_+$, $L \in \mathbb{R}^{n\times q}$, $H \in \mathbb{R}^{q\times q}$, and $\sigma \in \mathbb{R}_{>0}$. The satisfaction of (6.32) is equivalent to the feasibility of

$$\begin{bmatrix} \mathrm{He}(S_1^X) & S_2^X + \hat{P} \\ \bullet & N_1 + \mathrm{He}(S_3^X) \end{bmatrix} < \mathbf{0}, \quad \begin{bmatrix} \mathrm{He}(S_1^Y) & S_2^Y + \hat{P}_{T_2} \\ \bullet & N_2 + \mathrm{He}(S_3^Y) \end{bmatrix} < \mathbf{0} \tag{6.42a}$$

with respect to $X_1, Y_1, X_3, Y_3 \in \mathbb{R}^{n\times n}$, $X_2, Y_2 \in \mathbb{R}^{n\times q}$, $X_4, Y_4, X_6, Y_6 \in \mathbb{R}^{q\times n}$, $X_5, Y_5 \in \mathbb{R}^{q\times q}$, where:

$$\begin{aligned} \hat{P} &= \mathrm{diag}\{P_1, P_2\} \\ \hat{P}_{T_2} &= \mathrm{diag}\{P_1, P_2 e^{\sigma T_2}\} \\ N_1 &= \mathrm{diag}\{\mathbf{0}, -\sigma P_2\} \\ N_2 &= \mathrm{diag}\{\mathbf{0}, -\sigma e^{\sigma T_2} P_2\} \end{aligned} \tag{6.42b}$$

$$S_1^X = \begin{bmatrix} -X_1 + M^\mathsf{T} X_4 & -X_2 + M^\mathsf{T} X_5 \\ -X_4 & -X_5 \end{bmatrix} S_1^Y = \begin{bmatrix} -Y_1 + M^\mathsf{T} Y_4 & -Y_2 + M^\mathsf{T} Y_5 \\ -Y_4 & -Y_5 \end{bmatrix} \tag{6.42c}$$

$$S_3^X = \begin{bmatrix} (A-LM)^\mathsf{T} X_3 - M^\mathsf{T} H^\mathsf{T} X_6 & \mathbf{0} \\ L^\mathsf{T} X_3 + H^\mathsf{T} X_6 & \mathbf{0} \end{bmatrix} \quad S_3^Y = \begin{bmatrix} (A-LM)^\mathsf{T} Y_3 - M^\mathsf{T} H^\mathsf{T} Y_6 & \mathbf{0} \\ L^\mathsf{T} Y_3 + H^\mathsf{T} Y_6 & \mathbf{0} \end{bmatrix} \tag{6.42d}$$

$$\begin{aligned} S_2^Y &= \begin{bmatrix} Y_1^\mathsf{T}(A-LM) - Y_4^\mathsf{T} HM - Y_3 + M^\mathsf{T} Y_6 & Y_1^\mathsf{T} L + Y_4^\mathsf{T} H \\ Y_2^\mathsf{T}(A-LM) - Y_5^\mathsf{T} HM - Y_6 & Y_2^\mathsf{T} L + Y_5^\mathsf{T} H \end{bmatrix} \\ S_2^X &= \begin{bmatrix} X_1^\mathsf{T}(A-LM) - X_4^\mathsf{T} HM - X_3 + M^\mathsf{T} X_6 & X_1^\mathsf{T} L + X_4^\mathsf{T} H \\ X_2^\mathsf{T}(A-LM) - X_5^\mathsf{T} HM - X_6 & X_2^\mathsf{T} L + X_5^\mathsf{T} H \end{bmatrix} \end{aligned} \tag{6.42e}$$

Proof. First of all, notice that by defining the matrices

$$\mathcal{B} = \begin{bmatrix} \mathcal{F} \\ \mathrm{I} \end{bmatrix}, \mathcal{N}_1 = \begin{bmatrix} \mathbf{0} & \hat{P} \\ \bullet & N_1 \end{bmatrix}, \mathcal{N}_2 = \begin{bmatrix} \mathbf{0} & \hat{P}_{T_2} \\ \bullet & N_2 \end{bmatrix}$$

matrices $\mathcal{M}_1$ and $\mathcal{M}_2$ in (6.32) can be equivalently rewritten respectively as follows

$$\mathcal{M}_1 = \mathcal{B}^\mathsf{T}\mathcal{N}_1\mathcal{B}, \quad \mathcal{M}_2 = \mathcal{B}^\mathsf{T}\mathcal{N}_2\mathcal{B}. \tag{6.43}$$

Moreover, by defining

$$U = \begin{bmatrix} \mathbf{0}_{(2n+q)\times q} \\ \mathrm{I}_q \end{bmatrix}$$

the positive definiteness of P_2 is equivalent to the satisfaction of the following relations

$$\begin{aligned} U^\mathsf{T}\mathcal{N}_1 U &< \mathbf{0} \\ U^\mathsf{T}\mathcal{N}_2 U &< \mathbf{0}. \end{aligned} \tag{6.44}$$

Then by the projection lemma; see [99], (6.32) is verified if and only if there exist two matrices X, Y such that

$$\begin{aligned} \mathcal{N}_1 + \mathcal{B}^{\mathsf{T}\perp}_{r} X U_r^\perp + U^{\mathsf{T}\perp}_{r} X^\mathsf{T}\mathcal{B}_r^\perp &< \mathbf{0} \\ \mathcal{N}_2 + \mathcal{B}^{\mathsf{T}\perp}_{r} Y U_r^\perp + U^{\mathsf{T}\perp}_{r} Y^\mathsf{T}\mathcal{B}_r^\perp &< \mathbf{0} \end{aligned} \tag{6.45}$$

where $\mathcal{B}_r^\perp$ and $U_r^\perp$ are some matrices having as rows a basis of the row-null space respectively of $\mathcal{B}$ and U. Specifically, notice that in view of Fact 6.1, one can consider the following choice

$$\mathcal{B}_r^\perp = \begin{bmatrix} -\mathcal{F}_l^{-1} & \mathcal{F}_r \end{bmatrix} = \left[\begin{array}{cc|cc} -\mathrm{I} & \mathbf{0} & A - LM & L \\ M & -\mathrm{I} & -HM & H \end{array}\right]$$

while $U_r^\perp = \begin{bmatrix} \mathrm{I}_{2n+q} & \mathbf{0}_{(2n+q)\times q} \end{bmatrix}$. Thus, according to the following partitioning

$$X = \begin{bmatrix} X_1 & X_2 & X_3 \\ X_4 & X_5 & X_6 \end{bmatrix}, \quad Y = \begin{bmatrix} Y_1 & Y_2 & Y_3 \\ Y_4 & Y_5 & Y_6 \end{bmatrix}$$

relations (6.45) turn in (6.42a) and this concludes the proof. ■

The above result yields an equivalent condition to (6.32), that can be exploited to derive an efficient design procedure for the proposed observer, though introducing some conservatism. To this end, one needs to suitably manipulate (6.42a) in order to obtain conditions that are linear in the decision variables. Specifically, the two results given in the next sections provide two possible approaches to derive convex design procedures for the proposed observer.

Zero-order Sample-and-hold Intersample Scheme

Proposition 6.4. *If there exist $P_1 \in \mathcal{S}_+^n, P_2 \in \mathcal{S}_+^q$, a positive real scalar σ, a nonsingular matrix $X \in \mathbb{R}^{n\times n}$, and matrices $X_4, Y_4, X_6, Y_6 \in \mathbb{R}^{q\times n}, X_5, Y_5 \in \mathbb{R}^{q\times q}, J \in \mathbb{R}^{n\times q}$ such that*

$$\begin{bmatrix} \mathrm{He}(Q_1) & Q_2 + \hat{P} \\ \bullet & \mathrm{He}(Q_3) + N_1 \end{bmatrix} < \mathbf{0} \quad \begin{bmatrix} \mathrm{He}(R_1) & R_2 + \hat{P}_{T_2} \\ \bullet & \mathrm{He}(Q_3) + N_2 \end{bmatrix} < \mathbf{0} \tag{6.46}$$

where $\hat{P}, \hat{P}_{T_2}, N_1, N_2$ *are defined in* (6.42b) *and*

$$Q_1 = \begin{bmatrix} -X + M^\mathsf{T} X_4 & M^\mathsf{T} X_5 \\ -X_4 & -X_5 \end{bmatrix} \qquad R_1 = \begin{bmatrix} -X + M^\mathsf{T} Y_4 & M^\mathsf{T} Y_5 \\ -Y_4 & -Y_5 \end{bmatrix}$$

$$Q_2 = \begin{bmatrix} -X + M^\mathsf{T} X_6 + X^\mathsf{T} A - JM & J \\ -X_6 & \mathbf{0} \end{bmatrix} \quad R_2 = \begin{bmatrix} -X + M^\mathsf{T} Y_6 + X^\mathsf{T} A - JM & J \\ -Y_6 & \mathbf{0} \end{bmatrix}$$

$$Q_3 = \begin{bmatrix} A^\mathsf{T} X - M^\mathsf{T} J^\mathsf{T} & \mathbf{0} \\ J^\mathsf{T} & \mathbf{0} \end{bmatrix}$$

then $L = X^{-\mathsf{T}} J$ *and* $H = \mathbf{0}$ *are a solution to Problem 6.1.*

Proof. By selecting in (6.42a) $H = \mathbf{0}, X_1 = X_3 = Y_1 = Y_3 = X, X_2 = Y_2 = \mathbf{0}, X^\mathsf{T} L = J$ gives (6.46). Thus, thanks to Corollary 6.2 the result is proven. ■

It should be noticed that the above design procedure leads to the well known zero-order sample-and-hold scheme; see Figure 6.2.

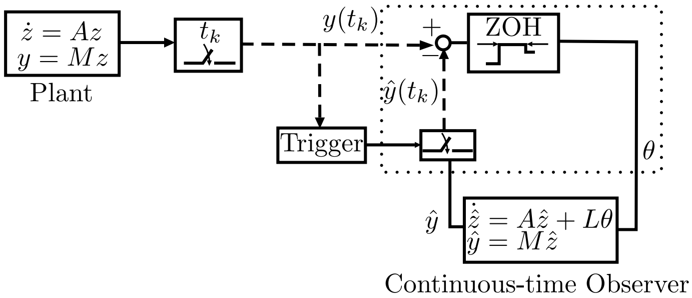


Figure 6.2: Zero-order sample-and-hold scheme

A Novel Observer Scheme

Proposition 6.5. *If there exist* $P_1 \in \mathcal{S}_+^n, P_2 \in \mathcal{S}_+^q$*, a positive real scalar* σ*, matrices* $X \in \mathbb{R}^{n\times n}, U, W \in \mathbb{R}^{q\times q}, J \in \mathbb{R}^{n\times q}$ *such that*

$$\begin{bmatrix} \mathrm{He}(Z_1) & Z_2 + \hat{P} \\ \bullet & Z_3 + N_1 \end{bmatrix} < \mathbf{0} \quad \begin{bmatrix} \mathrm{He}(Z_1) & Z_2 + \hat{P}_{T_2} \\ \bullet & \mathrm{He}(Z_3) + N_2 \end{bmatrix} < \mathbf{0} \tag{6.47}$$

where $\hat{P}, \hat{P}_{T_2}, N_1, N_2$ *are defined in* (6.42b) *and*

$$Z_1 = \begin{bmatrix} -X & U \\ \mathbf{0} & -U \end{bmatrix} \qquad Z_2 = \begin{bmatrix} -X + X^\mathsf{T} A - JM & J \\ -WM & W \end{bmatrix}$$

$$Z_3 = \begin{bmatrix} A^\mathsf{T} X - M^\mathsf{T} J^\mathsf{T} & \mathbf{0} \\ J^\mathsf{T} & \mathbf{0} \end{bmatrix}$$

then $L = X^{-\top} J$ *and* $H = U^{-\top} W$ *are a solution to Problem 6.1.*

Proof. By selecting in (6.42a) $X_1 = X_3 = Y_1 = Y_3 = X, X_2 = Y_2 = \mathbf{0}, X_4 = Y_4 = \mathbf{0}, X_6 = Y_6 = \mathbf{0}, X_5 = Y_5 = U, X^\top L = J, U^\top H = W$ gives (6.47). Thus, thanks to Corollary 6.2 the result is proven. ■

Remark 6.3. The above result gives rise to a novel observer scheme. Indeed, as a difference to Proposition 6.3 and Proposition 6.4, Proposition 6.5 does not impose any structural constraint on the gain H. This is a worthwhile novelty introduced by our approach with respect to classical approaches as [73, 104] and alike, where the choice of the gain H is *a priori* constrained. Thus, in general, the use of Proposition 6.5 may lead to observation schemes that are not encompassed either by Proposition 6.3 and Proposition 6.4 or by existing approaches.

Remark 6.4. The derivations of the design presented in Proposition 6.4 Proposition 6.5 consist in some particular choices of the slack variables X and Y introduced in Corollary 6.2. Therefore, when one is interested in solving Problem 6.1 for the largest achievable value of T_2, the design procedures arising from Proposition 6.4 and Proposition 6.5 may lead to conservative results. To overcome this problem, one can envision a two-stage procedure. Indeed, whenever L, H, σ and T_2 are fixed, condition (6.32) is linear in the decision variables. Thus, once the observer has been designed via one of the proposed methodologies, by testing the feasibility of (6.32) with respect to P_1, P_2 over a selected grid for σ, one may enable to enlarge the maximum allowable sampling interval T_2 for the considered design.

6.6 Numerical Examples

Example 6.1. In this first example, we want to show the improvement provided by our methodology with respect to existing results. Specifically, consider the example in [72], which is defined by the following data:

$$A = \begin{bmatrix} 0 & 1 \\ -4 & 0 \end{bmatrix}, M = \begin{bmatrix} 1 & 0 \end{bmatrix}.$$

As pointed out earlier, by setting $H = -ML$ in (6.3), the observer proposed in this chapter corresponds to the one in [72, 73]. Therefore, for a given gain L, by following the above selection for H, Theorem 6.2 can be used to provide an estimate of the maximum allowable sampling interval T_2. Hence to compare with [72], we consider

$$L = \begin{bmatrix} 4 \\ 0 \end{bmatrix}.$$

In this case, it turns out that the conditions of Theorem 6.2 are feasible for T_2 up to 0.42. This bound is about 5.18 times less conservative than the one in [72] ($T_2 = 0.081$). That is our methodology leads to an improvement on the estimation of the maximum allowable sampling interval of about 418%. On the other hand, Corollary 6.1 can also be used to design a new gain L_s to tentatively enlarge the maximum allowable sampling interval, still for the scheme proposed in [72]. Specifically, it turns out that, whenever the observer gain is designed via Corollary 6.1, conditions in (6.40) are feasible for T_2 up to 0.496, that is an improvement of about 18% with respect to the design in [72] . The observer gains obtained for $T_2 = 0.496$ are

$$L_s = \begin{bmatrix} 0.351 \\ -2.29 \end{bmatrix}, H_s = -ML_s = -0.351.$$

Figure 6.3 and Figure 6.4 report, respectively, the evolution of the estimation error and of θ,

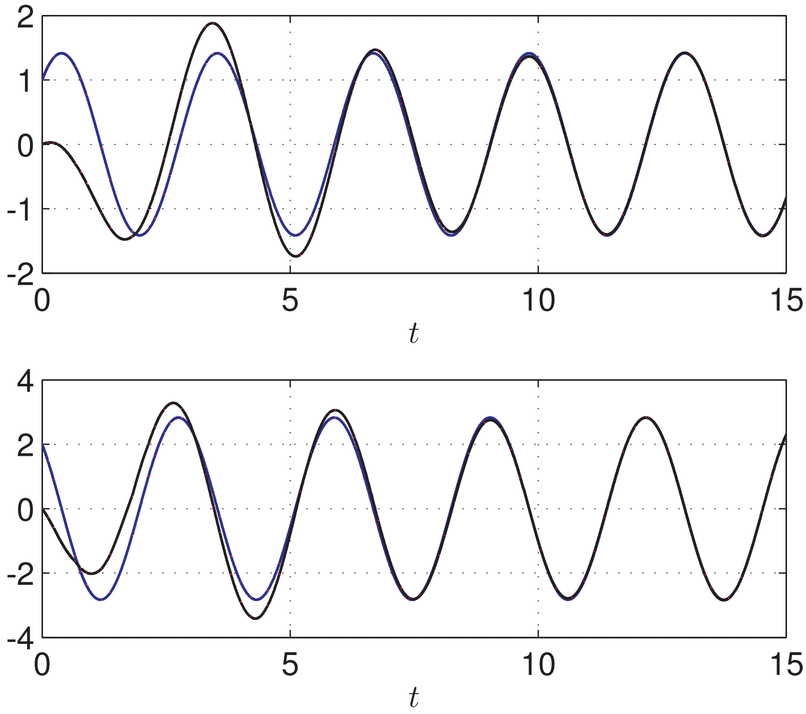


Figure 6.3: The evolution of the plant state z (blue) and of its estimate $\hat{z}$ (black) provided by the observer projected onto ordinary time. Above $z_1, \hat{z}_1$, below $z_2, \hat{z}_2$

both projected onto ordinary time. In this simulation $T_1 = 0.1$, and the sampling instances are chosen randomly according to a uniform distribution. Simulations show that the observer successfully reconstructs the plant state. Moreover, Figure 6.5 reports the evolution of the function V used in the proof of Theorem 6.2 projected onto ordinary time. Simulations show that the function V decreases during flows, and at jumps it is nonincreasing (in fact in this simulation it appears even decreasing).

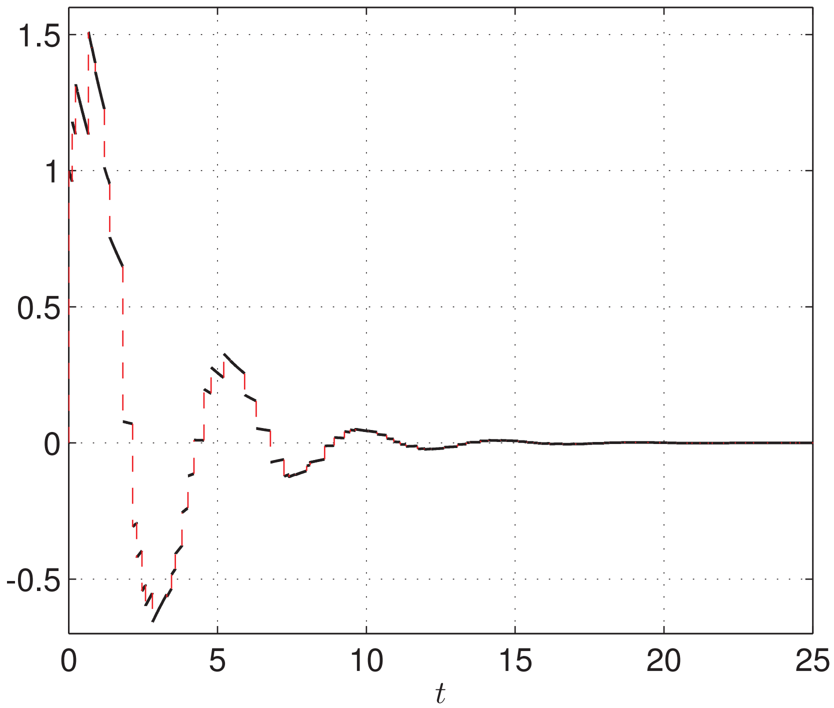


Figure 6.4: The evolution of θ projected onto ordinary time.

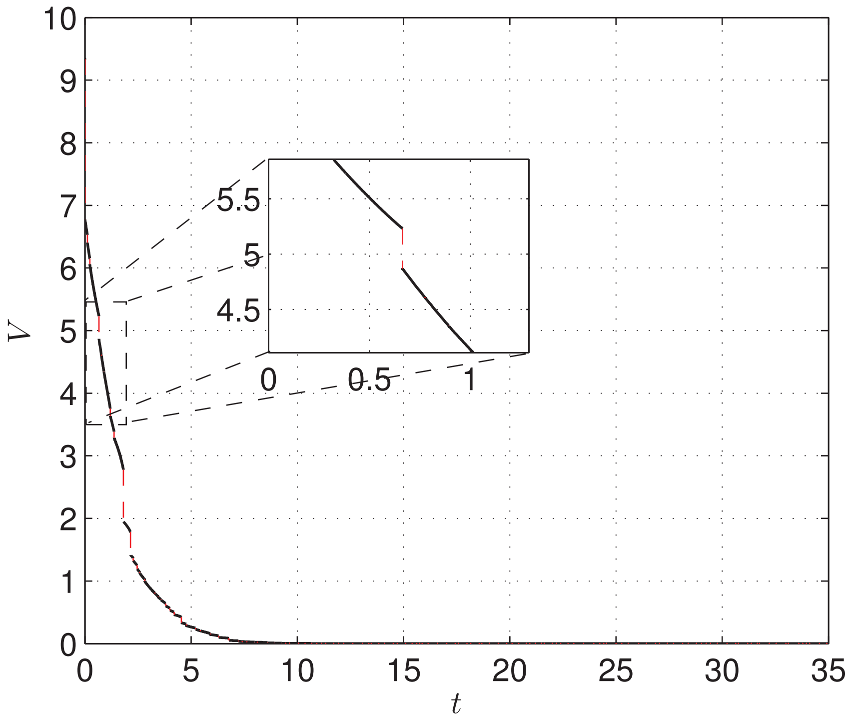


Figure 6.5: The evolution of the function V projected onto ordinary time.

Example 6.2. Consider the model of the longitudinal dynamics of the F8 aircraft in [71], whose state-space model is given by

$$\dot{x} = \begin{bmatrix} -0.8 & -0.006 & -12 & 0 \\ 0 & -0.014 & -16.6 & -32.2 \\ 1 & -10^{-4} & -1.5 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix} x$$
$$y = \begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & -1 & 1 \end{bmatrix} x.$$

The two outputs are respectively the pitch angle and the flight path angle. We want to design an observer for the considered plant while enlarging as much as possible the maximum transfer time T_2 allowable.

In Table 6.1, we report, for each design methodology, the values of the maximum T_2 for which conditions (6.32) are feasible along with the corresponding value of σ, and the two designed gains L and H. In each of these designs, the value of σ is selected so as to enlarge the value of T_2 ensuring the feasibility of the considered conditions. Concerning the design procedure derived by Proposition 6.4 and Proposition 6.5, as mentioned in Remark 6.4, to reduce as much as possible the conservatism in the estimate of the largest value of T_2 allowable, after a first design step, we performed a further analysis stage via Theorem 6.2. About the design procedure issued from Proposition 6.4, it is worthwhile to notice that, the

Design	σ	T_2	L	H
Proposition 6.3	0.6	4.7	$\begin{bmatrix} -0.712 & 0.872 \\ 1744 & -2133 \\ 2.69 & -3.28 \\ -8.13 & 9.95 \end{bmatrix}$	$\begin{bmatrix} 4.62 & -5.73 \\ 12.3 & -14.7 \end{bmatrix}$
Corollary 6.1	0.71	4.1	$\begin{bmatrix} 0.15702 & 0.42578 \\ -34.118 & -84.26 \\ 0.10341 & 0.24557 \\ 0.22093 & 0.53946 \end{bmatrix}$	$\begin{bmatrix} -0.221 & -0.539 \\ -0.118 & -0.294 \end{bmatrix}$
Proposition 6.4	0.97	3.54	$\begin{bmatrix} -0.216 & 0.216 \\ -21.6 & -36.1 \\ -0.00971 & -0.00372 \\ 0.1 & 0.134 \end{bmatrix}$	$\mathbf{0}_{2\times 2}$
Proposition 6.5	0.59	5.73	$\begin{bmatrix} -0.044 & 0.102 \\ -31.8 & -47.0 \\ 0.0184 & 0.0143 \\ 0.15 & 0.199 \end{bmatrix}$	$\begin{bmatrix} -0.258 & -0.0121 \\ -0.0172 & -0.236 \end{bmatrix}$

Table 6.1: Values of T_2 and σ and the designed observer gains L and H for the considered design procedures.

design conditions for the same observer scheme given in [104], (when they are specialized to the linear systems case), are feasible for T_2 up to 0.4. Namely, the proposed design, in this specific case, enables to enlarge the maximum allowable sampling interval of 8.85 times with respect to [104]. Moreover, it turns out that the design procedure issued from Proposition 6.5, in this specific case, provides the largest allowable value for T_2.

6.7 Comments and Conclusion

Building from the general ideas in [73], in this chapter we proposed a novel methodology to design, via linear matrix inequalities, an observer with intersample injection to exponentially estimate the state of a continuous-time linear system in the presence of sporadically available measurements. Specifically, pursuing a unified approach, we provided four design methodologies to design the observer, which are computationally efficient, *i.e.*, the design algorithm entails a time of computation which is polynomial with respect to the dimension of the data. Two of them lead back respectively to the observer scheme proposed in [73] and to the zero-order sample-and-hold proposed in [104], while the remaining lead to two completely novel schemes. Notice that, although we recover some existing schemes, the design procedures we propose are novel and, in some cases, outperform the corresponding existing design techniques, whenever they exist. To the best author knowledge, a unified approach for the systematic design of the class of observer presented in this chapter, which encompasses the observer in [73], ensuring exponential state estimation for a given value of the maximum sampling interval, has been presented for the first time in [41]. Furthermore, we would like to emphasize that, although this chapter is devoted to LTI plants, differently from Chapter 5, the extension to a wider class of plants, as the one considered in [104], is almost direct.

Concerning the possibility of adopting alternative frameworks to address the problem illustrated in this chapter, we would like to emphasize that employing a discrete-time approach, as the one in [29], would hardly lead to a tractable design for the proposed observer. In particular, notice that by discretizing the $(\varepsilon, \tilde{\theta})$ dynamics in between jumps would give rise to a discrete-time model for which the two gains L and H appear via a matrix exponential term, preventing from deriving a tractable design procedure via polytopic embedding strategies, as done in Chapter 5.

The proposed observer allows to provide an alternative solution to the state estimation problem in the presence of sporadic measurements with respect to the one proposed in Chapter 5. Moreover, the design of the observer proposed in this chapter appears simpler than the one in Chapter 5. In particular, we recall that in the design presented in Chapter 5, the total number of lines in the considered matrix inequalities is proportional to 2^n. Then, the complexity of the resulting design increases exponentially with the size of the plant. Instead, for the designs presented in this chapter, the number of lines and the number of scalar variables entailed by the resulting LMI feasibility problem increase polynomially in n;

Table 6.2 reports precisely these data for the mentioned designs. Hence, whenever the plant size increases, such designs are expected to be less complex from a numerical standpoint than the design in Chapter 5. However, the two considered approaches are deeply different and

Design	Number of scalar variables	Number of lines
Corollary 5.1	$(n+1)/2n + n^2 + nq$	$3n2^n$
Proposition 6.3	$(n+1)/2n + nq + q^2 + q(q+1)/2$	$2(n+q)$
Corollary 6.1	$(n+1)/2n + nq + q(q+1)/2$	$2(n+q)$
Proposition 6.4	$(n+1)/2n + q(q+1)/2 + n^2 + 4qn + 2q^2 + nq$	$2(n+q)$
Proposition 6.5	$(n+1)/2n + q(q+1)/2 + n^2 + 2q^2 + nq$	$2(n+q)$

Table 6.2: Number of lines and number of scalar variables entailed by the different designs.

both manifest advantages and disadvantages that prevent from overlooking one of the two solutions. In particular, the observer in Chapter 5 has been shown to be ISS with respect to measurement noise. So far, we did not succeed in showing such a property for the observer considered in this chapter. First investigations allowed to show that the observer considered in this chapter is finite-gain $\mathcal{L}_{2,\infty}$ stable from the measurement noise to the estimation error ε (see [95, Definition 3] for more details on this notion of stability for general hybrid systems with inputs and outputs). Obviously this latter property is weaker than the ISS proven for the observer in Chapter 5.

Another interesting point concerns the fact that the conditions worked out in this chapter to design the considered observer do not depend on the value of T_1, which is not the case for the design in Chapter 5. This observation gives rise to some important considerations. Among them, let us remark that in the case of periodic sampling, *i.e.*, $T_1 = T_2 = T$ the observer in Chapter 5 can be always designed via the proposed apparatus, provided that the pair (e^{AT}, M) is detectable, the same is not true for the observer with flow injection presented in this chapter. Specifically, observe that periodic sampling does not originate any change in the conditions considered within this chapter. The reason behind this matter stems from the fact that the observer presented in this chapter is designed by ensuring the decrease of a certain Lyapunov-like function within the flow set, whereas the behavior within the jump set, wherein T_1 comes into play, is rigidly prescribed by the structure of the considered observer. This remark fosters to consider more general jump maps for the observer presented in this chapter.

Still concerning the observer in Chapter 5, numerical experiments show that in general such an observer, thanks to the recommended design, allows to ensure larger values of the maximum allowable sampling interval, with respect to the schemes considered in this chapter. This gap between the two proposed observation schemes originates from the innate nature of such schemes. Indeed, the approach pursued in this chapter basically relies on the intrinsic robustness of the continuous-time observer used as a core to build up the considered observation scheme. Such a robustness explicitly appears in Proposition 6.1 and Proposition 6.2, respectively, in (ii) and in (b) in the form of certain bounds that can be fulfilled provided that T_2 is small enough. This discussion naturally establishes connections between our approach and the one in [73], which we recall is one of the inspiring approach leading to the ideas presented in this chapter. Completely different considerations hold for the scheme presented in Chapter 5. Indeed, such a scheme does not rely on any continuous-time observer. On the one hand, this fact does not suggest any strategy to derive first guidelines, as Proposition 6.1 and Proposition 6.2, for the design of such an observer. On the other hand, the fact of operating a reset of the whole estimate seems to better address the state estimation problem in the presence of sporadic measurements, at least in terms of maximum allowable sampling interval T_2. However, the use of such an observer barely allows to envision extensions to more involved settings of practical interest as the one considered for instance in [90] dealing with multi-outputs plants with asynchronous sporadic measurements. To give an hint of the difficulties encountered in this situation, here below we briefly illustrate the problem to solve in such a case and first attempts towards its solution.

Let us considers continuous-time linear time-invariant system in the form

$$\begin{aligned} \dot{z} &= Az \\ y_i &= M_i z \qquad \forall i = 1, 2, \ldots, p \end{aligned} \tag{6.48}$$

where $z \in \mathbb{R}^n$ and $y = (y_1, y_2, \ldots, y_p) \in \mathbb{R}^q$ are, respectively, the state and the measured output of the system, while A and M_i are constant matrices of appropriate dimensions. The goal is to design an observer providing an asymptotic estimate $\hat{z}$ of the state z whenever each of the component y_i of the vector y is available only at some time instances $t_k^{(i)}$, $k \in \mathbb{N}$, not known *a priori*. Obviously, whenever for each $k \in \mathbb{N}$, $t_k^{(1)} = t_k^{(2)} = \cdots = t_k^{(p)}$, one falls inside the focus of this chapter. However, whether this assumption does not hold, a modification of the scheme in (6.3) is needed. In particular, inspired by [100], we consider the following observer

$$\left\{ \begin{array}{ll} \left. \begin{array}{rcl} \dot{\hat{z}}(t) &=& A\hat{z}(t) + L\theta(t) \\ \dot{\theta}(t) &=& H\theta(t) \end{array} \right\} & \text{when } t \notin \{t_k^{(i)} : i = 1, 2, \ldots, p\}, k \in \mathbb{N} \\ & \\ \left. \begin{array}{rcl} \hat{z}(t^+) &=& \hat{z}(t) \\ \theta_i(t^+) &=& y_i(t) - M_i\hat{z}(t) \end{array} \right\} & \text{when } t = t_k^{(i)}, i \in \{1, 2, \ldots, p\}, k \in \mathbb{N} \end{array} \right. \tag{6.49}$$

In particular, in between measurements the above observer behaves as the one presented in this chapter. Instead, whenever a new measurement is received, only the corresponding components of the vector θ get updated via the received measurement. This proposed observer is currently part of our research activity. Specifically, a hybrid model of the considered observer have been constructed. First researches have shown how a generalization of the methodology presented in this chapter provides the right answer to tackle with the considered problem. In particular, the main point we addressed consists of reshaping Assumption 6.1 to match the asynchronous nature of the incoming measurements. Such a reshaping is inspired by the construction presented in [47]. Notice that, while the construction of such an observer is the natural extension of the one illustrated in this chapter, the design of an observer to tackle this problem within the framework considered in Chapter 5 does not appear clear.

The state estimation problem in the presence of asynchronous sporadic measurement is in part encompassed by the work in [100] dealing with state estimation of networked systems, though in [100] the authors focus on an emulation approach. However, differently from [100], we do not assume any scheduling behind the arrival of measurements. This enables to address a certain number of situations of practical interest, that are uncovered by the approaches building on protocols ; see, *e.g.*, [90] and the references therein.

The observers presented in this chapter and in the previous one can be used to build up controller architectures to asymptotically stabilize a linear plant in the presence of sporadic measurements. For this reason, to conclude this part of this dissertation, in the next chapter we present an observer-based controller, whose core is centered on the observer presented in Chapter 5. For brevity, we limit the analysis to a scheme built upon the observer in Chapter 5. Nevertheless, observe that the construction of a similar scheme building on the observer presented in this chapter can be considered without too much work.

7

OBSERVER-BASED CONTROL IN THE PRESENCE OF SPORADIC SENSING AND ACTUATION

“C’est par la logique qu’on démontre, c’est par l’intuition qu’on invente”.

– Henri Poincaré

7.1 Introduction

In this chapter, we consider the problem of stabilizing a linear time-invariant system in the presence of sporadic output measurements and sporadic access to the plant input. The plant is equipped with a zero-order hold device which stores the value of the input in between control input updates. We propose an observer-based controller consisting of a measurement-triggered observer, which experiences jumps in its state whenever a new measure is available, a state-feedback control law computed from the estimated state, and a copy of the zero-order hold device feeding the plant, which jumps whenever the control input is sent to the plant. The closed-loop system is modeled as a hybrid system that includes two timers triggering the two different events. The resulting hybrid system is analyzed as the cascade of hybrid systems and its asymptotic stability properties are established through a separation principle. In addition, a computationally design procedure based on LMIs is presented and illustrated in an example. First results pertaining to the problem presented in this chapter can be found in [43].

7.2 Problem Statement

7.2.1 System Description

Consider the following continuous-time linear system:

$$\mathcal{P}\colon \begin{cases} \dot{z} = Az + Bu \\ y = Mz \end{cases} \tag{7.1}$$

where $z \in \mathbb{R}^n$, $y \in \mathbb{R}^q$ and $u \in \mathbb{R}^p$ are, respectively, the state, the measured output, and the input of the system, while A, B and M are constant matrices of appropriate dimensions. Now, let us suppose that both the input channel and the output channel of system (7.1) are accessible in an intermittent fashion. Especially, assume the initial time $t_0 = 0$, let us assume that the output of system (7.1) is gathered only at time instances t_k, $k \in \mathbb{N}$, not known *a priori* and that the input channel grants its access only at time instances s_k, $k \in \mathbb{N}$, not known *a priori*. Analogously to the previous chapters, suppose that $\{t_k\}_{k=1}^{+\infty}$ and $\{s_k\}_{k=1}^{+\infty}$ are two strictly increasing unbounded real sequences of times and assume that there exist four positive real scalars $T_1^{\mathcal{O}} \leq T_2^{\mathcal{O}}, T_1^{\mathcal{U}} \leq T_2^{\mathcal{U}}$, such that

$$\begin{aligned} & T_1^{\mathcal{O}} \leq t_1 \leq T_2^{\mathcal{O}} \\ & T_1^{\mathcal{O}} \leq t_{k+1} - t_k \leq T_2^{\mathcal{O}} \qquad \forall k \in \mathbb{N} \\ & T_1^{\mathcal{U}} \leq s_1 \leq T_2^{\mathcal{U}} \\ & T_1^{\mathcal{U}} \leq s_{k+1} - s_k \leq T_2^{\mathcal{U}} \qquad \forall k \in \mathbb{N}. \end{aligned} \tag{7.2}$$

The problem studied in this chapter consists of designing an observer-based controller that

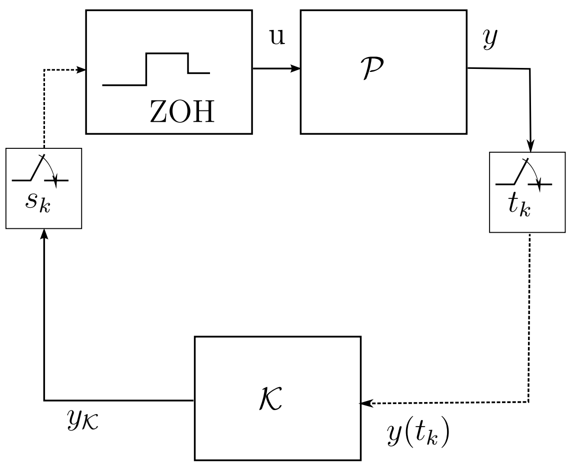


Figure 7.1: Continuous-time plant $\mathcal{P}$ controlled by the controller $\mathcal{K}$, which has intermittent access to the input channel and sporadic available measurements of the output y.

asymptotically stabilizes the resulting closed-loop system for any given sequences satisfying (7.2) providing measurements of the plant output and input access respectively.

Assuming that the arrival of a new measurement can be instantaneously detected by the controller, and that the controller is aware when a new sample is sent to the plant (such assumptions are not much severe and they can be fulfilled in real engineered systems; see, *e.g.*, [62, 113]) motivated by Chapter 5, we design an observer-based controller with jumps in its state $(\hat{z}, \hat{u})$, given by

$$\mathcal{K}\begin{cases} \left.\begin{array}{rcl} \dot{\hat{z}}(t) & = & A\hat{z}(t) + B\hat{u}(t) \\ \dot{\hat{u}}(t) & = & 0 \end{array}\right\} & \text{when } t \notin \{s_k\}_{k=1}^{+\infty} \cup \{t_k\}_{k=1}^{+\infty} \\ \hat{u}(t^+) \;=\; K\hat{z}(t) & \text{when } t \in \{s_k\}_{k=1}^{+\infty} \\ \hat{z}(t^+) \;=\; \hat{z}(t) + LM(z(t) - \hat{z}(t)) & \text{when } t \in \{t_k\}_{k=1}^{+\infty} \\ y_{\mathcal{K}}(t) \;=\; K\hat{z}(t) & \end{cases} \tag{7.3}$$

where L and K are two matrices of appropriate dimensions to be designed. The variable $\hat{z}$ represents the estimated state of the plant generated by the observer by means of the measured plant output y, while $\hat{u}$ stores the last value of the control input sent to the plant. Indeed, whenever a new sample of the control value is sent to the plant, the controller accordingly updates its internal variable $\hat{u}$ so as to memorize the signal applied to the plant input u. Furthermore, the plant is equipped with an event-based zero-order hold device, whose driving events are generated by new control input arriving. In particular, such a device stores the value of the last received input between two updates and it gets updated whenever a new control input is sent by the controller, see Figure 7.1. Thus, the input injected into the plant is piecewise constant, and specifically, for every integer $k \in \mathbb{N}$, $u(t) = K\hat{z}(s_k)$ for $t \in [s_k, s_{k+1})$, while $u(t) = u(0)$ for $t \in [0, s_1)$, where $u(0)$ denotes the initial condition of the zero-order hold device, which can be chosen arbitrarily. Moreover, notice that if $t \in \{s_k\}_{k=1}^{\infty} \cap \{t_k\}_{k=1}^{\infty}$ then both $\hat{z}$ and $\hat{u}$ are updated.

7.2.2 Hybrid Modeling

The fact that the closed-loop system experiences jumps when a new measurement is available or when the input channel grants access to the controller suggests that the dynamics of the closed-loop system can be described via a hybrid system. We provide a hybrid model that captures not only the behavior due to a single pair of sequences $\{t_k\}_{k=1}^{\infty}$, $\{s_k\}_{k=1}^{\infty}$, but each possible evolution generated by any sequence satisfying (7.2) respectively. This is a unique approach that, while leads to nonunique solutions, allows to establish a strong result for a family of sequences t_k and s_k.

The proposed modeling approach requires to model the time-driven mechanism governing the availability of measurements or of access to the plant input. To this end, as in Chapter 5, we add two auxiliary timer variables τ_1 and τ_2 to keep track of the duration of flows and to trigger jumps according to the mechanism in (7.3). In particular, this modeling procedure leads to a model that can be efficiently represented by the framework for hybrid systems proposed in [56]. To accomplish that, we make τ_1 and τ_2 decrease as ordinary time t increases and, whenever $\tau_1 = 0$ or $\tau_2 = 0$, reset it to any point in $[T_1^{\mathcal{U}}, T_2^{\mathcal{U}}]$ or $[T_1^{\mathcal{O}}, T_2^{\mathcal{O}}]$ respectively, so as to enforce (7.2). Then, after a jump occurs, the two timers are reset according to the following jump rule[1]:

$$\begin{bmatrix}\tau_1^+ \\ \tau_2^+\end{bmatrix} \in \begin{cases} \begin{bmatrix}[T_1^{\mathcal{U}}, T_2^{\mathcal{U}}] \\ \tau_2\end{bmatrix} & \text{if } \tau_1 = 0, \tau_2 \neq 0 \\ \begin{bmatrix}\tau_1 \\ [T_1^{\mathcal{O}}, T_2^{\mathcal{O}}]\end{bmatrix} & \text{if } \tau_1 \neq 0, \tau_2 = 0 \\ \left\{\begin{bmatrix}[T_1^{\mathcal{U}}, T_2^{\mathcal{U}}] \\ \tau_2\end{bmatrix}, \begin{bmatrix}\tau_1 \\ [T_1^{\mathcal{O}}, T_2^{\mathcal{O}}]\end{bmatrix}\right\} & \text{if } \tau_1 = \tau_2 = 0. \end{cases} \tag{7.4}$$

To capture this mechanism, we define a hybrid system $\mathcal{H}_c$ within the framework in [56]. In particular, take as a vector state $\tilde{x} = (z, u, \tau_1, \hat{z}, \hat{u}, \tau_2)$, and for each $x \in C = \mathbb{R}^n \times \mathbb{R}^p \times [0, T_2^{\mathcal{U}}] \times \mathbb{R}^n \times \mathbb{R}^p \times [0, T_2^{\mathcal{O}}]$ define the flow map as

$$F(x) \coloneqq \begin{bmatrix} Az + Bu \\ 0 \\ -1 \\ A\hat{z} + B\hat{u} \\ 0 \\ -1 \end{bmatrix}.$$

For each $x \in D$, define the jump map as

$$G(x) \coloneqq \begin{cases} G_1(x) & \text{if } x \in D_1 \setminus D_2 \\ G_2(x) & \text{if } x \in D_2 \setminus D_1 \\ \{G_1(x), G_2(x)\} & \text{if } x \in D_1 \cap D_2 \end{cases}$$

where for each $x \in D = D_1 \cup D_2$,

$$G_1(x) = \begin{bmatrix} z \\ K\hat{z} \\ [T_1^{\mathcal{U}}, T_2^{\mathcal{U}}] \\ \hat{z} \\ K\hat{z} \\ \tau_2 \end{bmatrix}, G_2(x) = \begin{bmatrix} z \\ u \\ \tau_1 \\ \hat{z} + LM(z - \hat{z}) \\ \hat{u} \\ [T_1^{\mathcal{O}}, T_2^{\mathcal{O}}] \end{bmatrix} \tag{7.5}$$

[1] The reason behind the choice considered to update the two timers in the case $\tau_1 = \tau_2 = 0$ will appear clear later.

$$\begin{aligned} D_1 &= \mathbb{R}^n \times \mathbb{R}^p \times \{0\} \times \mathbb{R}^n \times \mathbb{R}^p \times [0, T_2^{\mathcal{O}}] \\ D_2 &= \mathbb{R}^n \times \mathbb{R}^p \times [0, T_2^{\mathcal{U}}] \times \mathbb{R}^n \times \mathbb{R}^p \times \{0\}. \end{aligned} \tag{7.6}$$

These objects define a hybrid system $\mathcal{H}_c = (C, F, D, G)$ that represents the dynamics of the closed-loop system. Now, for the purpose of stabilization, consider the following invertible change of coordinates:

$$(z, u, \tau_1, \varepsilon, \tilde{u}, \tau_2) = (z, u, \tau_1, z - \hat{z}, u - \hat{u}, \tau_2) = x_e$$

which leads to the following model of the closed-loop system

$$\mathcal{H}_e \begin{cases} \dot{x}_e = F_e(x_e) & x_e \in C_e \\ x_e^+ \in G_e(x_e) & x_e \in D_e \end{cases} \tag{7.7a}$$

where $C_e = C, D_e = D_{1e} \cup D_{2e}, D_{1e} = D_1, D_{2e} = D_2$ and

$$F_e(x_e) = \begin{bmatrix} Az + Bu \\ 0 \\ -1 \\ A\varepsilon + B\tilde{u} \\ 0 \\ -1 \end{bmatrix}, G_{1e}(x_e) = \begin{bmatrix} z \\ K(z - \varepsilon) \\ [T_1^{\mathcal{U}}, T_2^{\mathcal{U}}] \\ \varepsilon \\ 0 \\ \tau_2 \end{bmatrix}, G_{2e}(x_e) = \begin{bmatrix} z \\ u \\ \tau_1 \\ (I - LM)\varepsilon \\ \tilde{u} \\ [T_1^{\mathcal{O}}, T_2^{\mathcal{O}}] \end{bmatrix} \tag{7.7b}$$

$$G_e(x_e) = \begin{cases} G_{1e}(x_e) & \text{if } x_e \in D_{1e} \setminus D_{2e} \\ G_{2e}(x_e) & \text{if } x_e \in D_{2e} \setminus D_{1e} \\ \{G_{1e}(x_e), G_{2e}(x_e)\} & \text{if } x_e \in D_{1e} \cap D_{2e}. \end{cases} \tag{7.7c}$$

Remark 7.1. Taking the union of the two reset laws whenever $\tau_1 = \tau_2 = 0$ in (7.4) ensures that the resulting jump map G_e is outer semicontinuous relatively to D. This fact can be proven by directly resorting to the definition of outer semicontinuity for set-valued mappings given in Appendix D. This fact, along with the continuity of the flow map, $D_e \subset \operatorname{dom} G_e$, $C_e \subset \operatorname{dom} F_e$, and the closedness of the sets C_e and D_e ensures that hybrid system (7.7) satisfies Assumption 4.1. This is a key property that will be used in the sequel. Observe that having a hybrid system satisfying Assumption 4.1 may not be trivial and it actually derives from suitable choices done throughout the modeling stage. Several cases of hybrid systems not matching Assumption 4.1 can be encountered in the literature; see, *e.g.*, the hysteretic quantizer in [22].

Concerning the existence of solutions to system $\mathcal{H}_e$, by relying on the concept of solution given in Definition 4.5, it is straightforward to check that every $\phi \in \mathcal{S}_{\mathcal{H}_e}(C_e \cup D_e)$ is complete. Moreover, the following properties hold:

- For every $(t, j) \in \operatorname{dom} \phi$ such that $(t, j + 1) \in \operatorname{dom} \phi$ and $\phi(t, j) \in D_{2e} \setminus D_{1e}$, one has $(t, j + 2) \notin \operatorname{dom} \phi$,

- For every $(t,j) \in \operatorname{dom}\phi$ such that $(t,j+1) \in \operatorname{dom}\phi$ and $\phi(t,j) \in D_{1e} \setminus D_{2e}$, one has $(t,j+2) \notin \operatorname{dom}\phi$,

- For every $(t,j) \in \operatorname{dom}\phi$ such that $(t,j+1) \in \operatorname{dom}\phi$ and $\phi(t,j) \in D_{1e} \cap D_{2e}$, we have either $\phi(t,j+1) \in D_{1e} \setminus D_{2e}$ or $\phi(t,j+1) \in D_{2e} \setminus D_{1e}$.

In other words, at most two jumps can occur consecutively without flowing. Furthermore, for every maximal solution ϕ to $\mathcal{H}_e$, due to (7.2), every $(t,j) \in \operatorname{dom}\phi$ such that $(t,s) \in \operatorname{dom}\phi$, for some $s \in \{j+1, j+2\}$, implies that $\left\{[t, t+\min\{T_1^{\mathcal{O}}, T_1^{\mathcal{U}}\}] \times \{s\}\right\} \subset \operatorname{dom}\phi$. Essentially, the domain of the solutions to $\mathcal{H}_e$ manifests an average dwell-time property, with dwell time $\tau_D = \min\{T_1^{\mathcal{O}}, T_1^{\mathcal{U}}\}$ and offset $N_0 = 2$; see, *e.g.*, [56, Example 2.15]. Such a property imposes a strictly positive uniform lower bound on the length of every flow interval, preventing from the existence of Zeno solutions.

Remark 7.2. A notable property enforced by timer τ_1 is that, for every maximal solution to (7.7), there exists $(T,J) \in \operatorname{dom}\phi$ satisfying $T + J \leq T_2^{\mathcal{U}} + 1$, such that $\phi(T,J) \in D_{1e}$, which implies that $\tilde{u}(T, J+1) = 0$. Then, since solutions to (7.7) cannot leave the set $\mathbb{R}^n \times \mathbb{R}^p \times [0, T_2^{\mathcal{U}}] \times \mathbb{R}^n \times \{0\} \times [0, T_2^{\mathcal{O}}]$, it follows that for every initial condition $\phi(0,0) \in C_e \cup D_e$, $\tilde{u}$ converges to zero in finite hybrid time. Moreover, notice that to make the hybrid system (7.7) an accurate description of the real time-triggered phenomenon, which governs the update process, τ_1 and τ_2 have to belong to the intervals $[0, T_2^{\mathcal{U}}]$ and $[0, T_2^{\mathcal{O}}]$ respectively, which is a property that is guaranteed by the definition of C_e and D_e.

In this chapter, we consider the following notions for a general hybrid system $\mathcal{H}$ with state in $\mathbb{R}^{\ell}$.

Definition 7.1. ([56, Definition 7.1.]) Let $\mathcal{A} \subset \mathbb{R}^{\ell}$ be a compact set. The set $\mathcal{A}$ is

- stable for $\mathcal{H}$ if for every $\epsilon > 0$ there exists $\delta > 0$ such that every solution to $\mathcal{H}$ with $|\phi(0,0)|_{\mathcal{A}} \leq \delta$ satisfies $|\phi(t,j)|_{\mathcal{A}} \leq \epsilon$ for all $(t,j) \in \operatorname{dom}\phi$;

- locally pre-attractive for $\mathcal{H}$ if there exists $\mu > 0$ such that every solution ϕ to $\mathcal{H}$ with $|\phi(0,0)|_{\mathcal{A}} \leq \mu$ is bounded and, if ϕ is complete, then also $\lim_{t+j \to +\infty} |\phi(t,j)|_{\mathcal{A}} = 0$;

- locally pre-asymptotically stable (LpAS) for $\mathcal{H}$, if it is both stable and locally pre-attractive for $\mathcal{H}$;

- globally pre-asymptotically stable (GpAS) for $\mathcal{H}$, if it is both stable and locally pre-attractive for $\mathcal{H}$ for every $\mu > 0$.

Definition 7.2. ([56]) A set $\mathcal{A} \subset \mathbb{R}^{\ell}$ is strongly forward pre-invariant for $\mathcal{H}$, if for every maximal solution ϕ to $\mathcal{H}$, $\operatorname{rge}\phi \subset \mathcal{A}$.

Remark 7.3. In referring to complete solutions, we will drop the term "pre" from the above definitions, which leads respectively to locally asymptotically stable (LAS), globally asymptotically stable (GAS), and strongly forward invariant.

Then, by introducing the set

$$\mathcal{A} = \{0\} \times \{0\} \times [0, T_2^{\mathcal{U}}] \times \{0\} \times \{0\} \times [0, T_2^{\mathcal{O}}] \tag{7.8}$$

for which, for every $x_e \in C_e \cup D_e \cup G_e(D_e)$, $|x_e|_{\mathcal{A}} = \|(z,\, u,\, \varepsilon,\, \tilde{u})\|$, the problem we solve is as follows:

Problem 7.1. Given the matrices A, B, and M of appropriate dimensions and four positive scalars $T_1^{\mathcal{U}} \leq T_2^{\mathcal{U}}$, $T_1^{\mathcal{O}} \leq T_2^{\mathcal{O}}$, design matrices $L \in \mathbb{R}^{n\times q}$ and $K \in \mathbb{R}^{p\times n}$ such that the set $\mathcal{A}$ in (7.8) is globally asymptotically stable for the hybrid system (7.7).

To cope with this problem, we treat (7.7) as the cascade of two hybrid systems (modulo the coupling effect, yet vanishing in finite hybrid-time, as shown in Remark 7.2, induced by $\tilde{u}$ on the ε dynamics). Namely, this cascade is composed by the ε dynamics along with its timer τ_2, which enters into the (z, u, τ_1) dynamics. By pursuing this approach, we are able to solve Problem 7.1 without the need of finding a Lyapunov function for the whole hybrid system (7.7), which appears as a nontrivial problem.

7.3 Main results

7.3.1 A solution via a Separation Principle

In this section, we provide a solution to Problem 7.1 that relies on the properties inherited from the components of the closed-loop system, namely, the observer and the controller subsystems. Specifically, let us consider the following assumptions.

Assumption 7.1 (Observer subsystem)**.** The hybrid system

$$\begin{cases} \left.\begin{array}{rcl} \dot{\varepsilon} & = & A\varepsilon \\ \dot{\tau}_2 & = & -1 \end{array}\right\} & (\varepsilon, \tau_2) \in C_o \\ \left.\begin{array}{rcl} \varepsilon^+ & = & (I - LM)\varepsilon \\ \tau_2^+ & \in & [T_1^{\mathcal{O}}, T_2^{\mathcal{O}}] \end{array}\right\} & (\varepsilon, \tau_2) \in D_o \end{cases} \tag{7.9a}$$

where

$$C_o = \mathbb{R}^n \times [0, T_2^{\mathcal{O}}],\ \ D_o = \mathbb{R}^n \times \{0\} \tag{7.9b}$$

has the set $\mathcal{A}_o = \{0\} \times [0, T_2^{\mathcal{O}}]$ GAS. △

Assumption 7.2 (Controller subsystem)**.** The hybrid system

$$\begin{cases} \left.\begin{array}{rcl} \dot{z} & = & Az + Bu \\ \dot{u} & = & 0 \\ \dot{\tau}_1 & = & -1 \end{array}\right\} & (z, u, \tau_1) \in C_{\mathcal{K}} \\ \\ \left.\begin{array}{rcl} z^+ & = & z \\ u^+ & = & Kz \\ \tau_1^+ & \in & [T_1^{\mathcal{U}}, T_2^{\mathcal{U}}] \end{array}\right\} & (z, u, \tau_1) \in D_{\mathcal{K}} \end{cases} \tag{7.10a}$$

where

$$C_{\mathcal{K}} = \mathbb{R}^n \times \mathbb{R}^p \times [0, T_2^{\mathcal{U}}],\ D_{\mathcal{K}} = \mathbb{R}^n \times \mathbb{R}^p \times \{0\} \tag{7.10b}$$

has the set $\mathcal{A}_{\mathcal{K}} = \{0\} \times \{0\} \times [0, T_2^{\mathcal{U}}]$ GAS. △

A sufficient condition guaranteeing that Assumption 7.1 holds is given in Chapter 5, while a sufficient condition for Assumption 7.2 to hold will be given in Proposition 7.1. The following result establishes GAS of the set $\mathcal{A}$ for the closed-loop system (7.7) under the two aforementioned assumptions. Before state such results, let us consider the following definition.

Definition 7.3. Given a hybrid systems $\mathcal{H} = (C, F, D, G)$ with state in $\mathbb{R}^n$, and let $O \subset \mathbb{R}^n$. We denote, $\mathcal{H}|_O = (C \cap O, F, D \cap O, G)$.

Remark 7.4. Namely, $\mathcal{H}|_O$ is the restriction of the dynamics of $\mathcal{H}$ to the set O. Notice that in the above definition, any property is required for O. In particular, $O \cap (C \cup D)$ could be empty leading to a restriction having no solutions.

Theorem 7.1. Let Assumption 7.1 and Assumption 7.2 hold. Then, the set $\mathcal{A}$ defined in (7.8) is GAS for system (7.7). □

The proof of this theorem is inspired by the idea in [122, Theorem 1]. Specifically, we base our proof on [56, Corollary 7.24], which requires the satisfaction of Assumption 4.1 (hybrid basic assumption on data), that is satisfied by (7.7). Since the proof is rather involved, for the sake of clarity, we firstly provide a list of the main steps carried out.

As a first step, to situate the analysis within the focus of [56, Corollary 7.24], which works with compact sets, we select an arbitrarily compact set J having $\mathcal{A}$ in its interior and we build the following auxiliary system $\mathcal{H}_{eJ} := \mathcal{H}_e|_J$. For such a system, we prove that the set $\mathcal{A}$ is GpAS, by performing the following steps:

(a) Prove that there exist two compact sets $\mathcal{J}_{\tilde{u}} \supset \mathcal{J}_{\varepsilon}$ and $\mathcal{J}_{\varepsilon} \supset \mathcal{A}$ such that:

  (a.1) $\mathcal{J}_{\tilde{u}}$ is GpAS for $\mathcal{H}_{eJ}$

  (a.2) $\mathcal{J}_{\varepsilon}$ is GpAS for $\mathcal{H}_{eJ}|_{\mathcal{J}_{\tilde{u}}}$

(b) Since $\mathcal{J}_{\varepsilon} \subset \mathcal{J}_{\tilde{u}}$, applying [56, Corollary 7.24] allows to conclude that $\mathcal{J}_{\varepsilon}$ is GpAS for $\mathcal{H}_{eJ}$

(c) Prove that $\mathcal{H}_{eJ}|_{\mathcal{J}_{\varepsilon}}$ has $\mathcal{A}$ GpAS

(d) Since $\mathcal{J}_{\varepsilon} \supset \mathcal{A}$, thanks to (c) applying [56, Corollary 7.24] allows to conclude that $\mathcal{A}$ is GpAS for $\mathcal{H}_{eJ}$.

Step (a) is performed by selecting a compact set $\mathcal{J}_{\tilde{u}}$ such that for every $x = (z, u, \tau_1, \varepsilon, \tilde{u}, \tau_2) \in \mathcal{J}_{\tilde{u}}$ $\tilde{u} = 0$, and compact set $\mathcal{J}_{\varepsilon} \subset \mathcal{J}_{\tilde{u}}$ such that for every $x = (z, u, \tau_1, \varepsilon, \tilde{u}, \tau_2) \in \mathcal{J}_{\varepsilon}$ is such that $\varepsilon = 0$. In particular, $(a.1)$ is established by using finite hybrid-time convergence of $\tilde{u}$ to zero. While, $(a.2)$ follows from Assumption 7.1. Finally, (c) follows from Assumption 7.2.

From GpAS of $\mathcal{A}$ for $\mathcal{H}_{eJ}$ and the fact that $\mathcal{A} \subset \mathrm{Int}J$, we prove LpAS of $\mathcal{A}$ for $\mathcal{H}_e$, which turns out to be LAS, due to completeness of the maximal solution to $\mathcal{H}_e$. Finally, from LAS and the fact that the compact set J can be selected arbitrarily large, GAS is established via homogeneity arguments.

The following two results basically establish the point (a) here above.

Claim 7.1. Define the closed set $\mathcal{A}_{\tilde{u}} = \mathbb{R}^n \times \mathbb{R}^p \times [0, T_2^{\mathcal{U}}] \times \mathbb{R}^n \times \{0\} \times [0, T_2^{\mathcal{O}}]$. Let Assumption 7.1 hold. Then, $\mathcal{H}_e|_{\mathcal{A}_{\tilde{u}}}$ has the closed set $\mathcal{A}_\varepsilon = \mathbb{R}^n \times \mathbb{R}^p \times [0, T_2^{\mathcal{U}}] \times \{0\} \times \{0\} \times [0, T_2^{\mathcal{O}}]$ GAS. □

The proof of the above claim is given in Appendix C.

Lemma 7.1. Pick any positive real scalars $M_z, M_u, M_\varepsilon, M_{\tilde{u}}, M_{\tau_1}, M_{\tau_2}$, and define the compact set $J = M_z\mathbb{B} \times M_{\tilde{u}} \times M_{\tau_1}\mathbb{B} \times M_{\tilde{u}}\mathbb{B} \times \mathbb{B}M_{\tau_2}$. Let $\mathcal{A}_\varepsilon$ be the set defined in Claim 7.1. Assumption 7.1 implies that hybrid system $\mathcal{H}_{eJ} := \mathcal{H}_e|_J$ has the compact set $\mathcal{J}_\varepsilon = (J \cup G_e(J)) \cap \mathcal{A}_\varepsilon$ GpAS. □

The proof of the above lemma is given later.

Now we are in position to provide the proof of Theorem 7.1.

Proof of Theorem 7.1. We first show that under Assumption 7.1 and Assumption 7.2 the set $\mathcal{A}$ defined in (7.8) is LAS for system (7.7) and that its basin of attraction contains every initial condition such that the resulting trajectory is bounded. Then we prove that LAS of $\mathcal{A}$ ensures that every maximal solution to (7.7) is bounded, allowing to extend the basin of attraction of $\mathcal{A}$ to include $C_e \cup D_e$, yielding GAS for $\mathcal{A}$.

Pick six arbitrarily large positive scalars $M_z, M_u, M_\varepsilon, M_{\tilde{u}}, M_{\tau_1}, M_{\tau_2}$ such that the compact set $J = M_z\mathbb{B} \times M_u\mathbb{B} \times M_{\tau_1}\mathbb{B} \times M_\varepsilon\mathbb{B} \times M_{\tilde{u}}\mathbb{B} \times M_{\tau_2}\mathbb{B}$ contains $\mathcal{A}$ in its interior. Define the closed set

$$\mathcal{A}_\varepsilon = \mathbb{R}^n \times \mathbb{R}^p \times [0, T_2^{\mathcal{U}}] \times \{0\} \times \{0\} \times [0, T_2^{\mathcal{O}}].$$

According to Lemma 7.1, which uses Assumption 7.1, the set $\mathcal{J}_\varepsilon = (J \cup G(J)) \cap \mathcal{A}_\varepsilon$ is GpAS for system $\mathcal{H}_{eJ} := \mathcal{H}_e|_J$. Moreover, thanks to Assumption 7.2 and by following the same steps as in the proof of Lemma 7.1, it turns out that the set $\mathcal{A}$ is GpAS for system $\mathcal{H}_{eJ}|_{\mathcal{A}_\varepsilon}$, and since $\mathcal{J}_\varepsilon \subset \mathcal{A}_\varepsilon$, similarly to [56, Proposition 3.32], it follows that $\mathcal{A}$ is GpAS for system $\mathcal{H}_{eJ}|_{\mathcal{J}_\varepsilon}$. Thus, since system $\mathcal{H}_{eJ}$ satisfies Assumption 4.1, $\mathcal{A}$ and $\mathcal{J}_\varepsilon$ are compact, and $\mathcal{A} \subset \mathcal{J}_\varepsilon$, thanks to [56, Corollary 7.24], the set $\mathcal{A}$ is LpAS for system $\mathcal{H}_{eJ}$ and its basin of attraction correspond to the one of $\mathcal{J}_\varepsilon$, which is equal to J establishing GpAS of $\mathcal{A}$ for $\mathcal{H}_{eJ}$.

Building on GpAS of $\mathcal{A}$ for $\mathcal{H}_{eJ}$, we establish LAS of the same set for $\mathcal{H}_e$. First we show that GpAS of $\mathcal{A}$ for $\mathcal{H}_{eJ}$ implies stability of $\mathcal{A}$ for $\mathcal{H}_e$.

To this end, pick $\epsilon > 0$ and suppose without loss of generality that $\mathcal{A} + \epsilon\mathbb{B} \subset \mathrm{Int}J$, such a choice is always possible due to $\mathcal{A} \subset \mathrm{Int}J$ by selecting ϵ small enough. From GpAS of $\mathcal{A}$ for system $\mathcal{H}_{eJ}$, it is always possible to pick $\delta > 0$ such that for every solution ϕ to $\mathcal{H}_{eJ}$, $|\phi(0,0)|_{\mathcal{A}} \leq \delta$ implies $|\phi(t,j)|_{\mathcal{A}} \leq \epsilon$ for all $(t,j) \in \operatorname{dom}\phi$. Now, from Lemma C.1, it follows that $\mathcal{S}_{\mathcal{H}_e}(\mathcal{A}+\delta\mathbb{B}) \subset \mathcal{S}_{\mathcal{H}_{eJ}}(\mathcal{A}+\delta\mathbb{B})$. Pick any $\psi \in \mathcal{S}_{\mathcal{H}_e}(\mathcal{A}+\delta\mathbb{B})$. Then thanks to the selection considered for δ one has for all $(t,j) \in \operatorname{dom}\psi$ that $|\psi(t,j)|_{\mathcal{A}} \leq \epsilon$. Hence, since the above arguments can be performed for any selection of $\epsilon > 0$, it follows that $\mathcal{A}$ is stable for $\mathcal{H}_J$.

Now we prove local attractivity of $\mathcal{A}$ for $\mathcal{H}_e$. The proof follows similar steps to the proof of stability here above. In particular, pick the same pair (ϵ, δ) from above. Since for such a pair we shown that $\phi \in \mathcal{S}_{\mathcal{H}_e}(\mathcal{A} + \delta\mathbb{B})$ implies that ϕ is a maximal solution to $\mathcal{H}_{Je}$ and maximal solutions to $\mathcal{H}_e$ are complete. Hence, from GpAS of $\mathcal{A}$ for $\mathcal{H}_{eJ}$, it follows that every $\phi \in \mathcal{S}_{\mathcal{H}_e}(\mathcal{A} + \delta\mathbb{B})$ converges to $\mathcal{A}$. Then, $\mathcal{A}$ is locally attractive for $\mathcal{H}_e$. This latter property along with the stability proven above establish that $\mathcal{A}$ is LAS for $\mathcal{H}_e$. Furthermore, since $M_z, M_u, M_\varepsilon, M_{\tilde{u}}, M_{\tau_1}, M_{\tau_2}$ can be selected arbitrarily large, for every maximal bounded solution ϕ to $\mathcal{H}_e$, there exists a suitable choice for $M_z, M_u, M_\varepsilon, M_{\tilde{u}}, M_{\tau_1}, M_{\tau_2}$ such that $\operatorname{rge}\phi \subset \operatorname{Int}J$. Hence, such a ϕ is a complete solution to $\mathcal{H}_{eJ}$ and it converges to $\mathcal{A}$ being $\mathcal{A}$ GpAS for $\mathcal{H}_e$. Thus, the basin of attraction of $\mathcal{A}$ contains each point from which maximal solutions to $\mathcal{H}_e$ are bounded. Thus to establish GAS, we show that every maximal solution to (7.7) is bounded, that is the basin of attraction of $\mathcal{A}$ includes $C_e \cup D_e$.

For each positive λ, define $M_\lambda = \operatorname{diag}(\lambda \mathrm{I}, \lambda \mathrm{I}, 1, \lambda \mathrm{I}, \lambda \mathrm{I}, 1)$, and notice that for each[2] $x \in C_e \cup D_e \cup G_e(D_e)$, one has $|M_\lambda x|_{\mathcal{A}} = \lambda \|(z, u, \varepsilon, \tilde{u})\|$. Pick any maximal solution ψ to (7.7) and denote

$$(t, j) \mapsto \psi(t, j) = (z(t, j), u(t, j), \tau_1(t, j), \varepsilon(t, j), \tilde{u}(t, j), \tau_2(t, j)).$$

From LAS of $\mathcal{A}$ for (7.7), there exists $\mu > 0$ such that every $\phi \in \mathcal{S}_{\mathcal{H}_e}(\mathcal{A} + \mu\mathbb{B})$ is bounded. Select a small enough $\lambda^\star > 0$ such that $\lambda^\star \|(z(0,0), u(0,0), \varepsilon(0,0), \tilde{u}(0,0))\| \leq \mu$, then

$$|M_\lambda^\star \psi(0,0)|_{\mathcal{A}} \leq \mu.$$

For each $(t, j) \in \operatorname{dom}\psi$, consider the function $(t, j) \mapsto M_\lambda^\star \psi(t, j)$, and notice that according to Lemma C.2 $M_\lambda^\star \psi$ is a maximal solution to (7.7). In particular, due to the selection considered for μ, one has $M_\lambda^\star \psi \in \mathcal{S}_{\mathcal{H}_e}(\mathcal{A} + \mu\mathbb{B})$, therefore $M_\lambda^\star \psi$ is bounded. Since the above arguments hold for any maximal solutions, boundedness of maximal solutions to (7.7) is established and this finishes the proof. ■

Now, the proof of Lemma 7.1 is given. Such a proof uses the definition of uniform pre-attractivity of a closed set for a general hybrid system $\mathcal{H}$ with state in $\mathbb{R}^\ell$.

Definition 7.4. ([56, Definition 6.24]) A compact set $\mathcal{A} \subset \mathbb{R}^\ell$ is said to be uniformly pre-attractive from a set $S \subset \mathbb{R}^\ell$ for $\mathcal{H}$ if every $\phi \in \mathcal{S}_{\mathcal{H}}(S)$ is bounded and for every $\epsilon > 0$ there exists $T > 0$ such that $|\phi(t, j)|_{\mathcal{A}} \leq \epsilon$ for every $\phi \in \mathcal{S}_{\mathcal{H}}(S)$ and $(t, j) \in \operatorname{dom}\phi$ with $t + j \geq T$.

Proof of Lemma 7.1. Let $\mathcal{A}_{\tilde{u}}$ be the set defined in Claim 7.1. The set $\mathcal{J}_{\tilde{u}} = (J \cup G_e(J)) \cap \mathcal{A}_{\tilde{u}}$ is compact, and uniformly pre-attractive for system $\mathcal{H}_{eJ}$ from any neighborhood of $\mathcal{J}_{\tilde{u}}$. In particular, notice that each solution to $\mathcal{H}_{eJ}$ is bounded and that for each complete solution ϕ to $\mathcal{H}_{eJ}$, as pointed out in Remark 7.2, there exist a (solution independent) strictly positive scalar T, such that $t + j \geq T$, with $(t, j) \operatorname{dom}\phi$, implies $\phi(t, j) \in \mathcal{A}_{\tilde{u}}$. This is enough to show uniform pre-attractivity of $\mathcal{J}_{\tilde{u}}$ for $\mathcal{H}_{eJ}$ from any neighborhood of $\mathcal{J}_{\tilde{u}}$. Then, since $\mathcal{H}_{eJ}$ satisfies Assumption 4.1, thanks to [56, Proposition 7.5], along with global (uniform) pre-attractivity of the set $\mathcal{J}_{\tilde{u}}$ for $\mathcal{H}_{eJ}$ shown right above, it follows that $\mathcal{J}_{\tilde{u}}$ is GpAS for $\mathcal{H}_{eJ}$.

[2]Notice that M_λ amounts to the nonproper standard dilation defined in [57, Definition 3.7]. In particular, $\mathcal{H}_e$ is homogeneous of degree zero with respect to the nonstandard dilation; see [123].

Now, consider the system $\mathcal{H}_{eJ}|_{\mathcal{J}_{\tilde{u}}} = \mathcal{H}_e|_{J\cap\mathcal{J}_{\tilde{u}}}$. Since $J\cap\mathcal{J}_{\tilde{u}} \subset \mathcal{A}_{\tilde{u}}$, by containment arguments (see; [56, Proposition 3.32]), it follows that every solution to $\mathcal{H}_{eJ}|_{\mathcal{J}_{\tilde{u}}}$ is a solution to $\mathcal{H}_{eJ}|_{\mathcal{A}_{\tilde{u}}}$. Thus, from Claim 7.1, the set $\mathcal{J}_\varepsilon$ is GpAS for $\mathcal{H}_{eJ}|_{\mathcal{J}_{\tilde{u}}}$. Furthermore, as $\mathcal{J}_\varepsilon \subset \mathcal{J}_{\tilde{u}}$, from [56, Corollary 7.24], it follows that $\mathcal{J}_\varepsilon$ is LpAS for $\mathcal{H}_{eJ}$ and its basin of pre-attraction coincides with the one of $\mathcal{J}_{\tilde{u}}$, which in turn coincides with J. Therefore $\mathcal{J}_\varepsilon$ is GpAS for $\mathcal{H}_{eJ}$. ■

7.3.2 Sufficient Conditions

Now, we provide sufficient conditions guaranteeing that the stated assumptions hold.

The observer gain L can be already designed to satisfy Assumption 7.1 via Corollary 5.1 on Page 157. To design the controller K ensuring that Assumption 7.2 is verified, as follows a constructive methodology is offered. Such a methodology basically uses ideas from [56, Example 3.21].

Proposition 7.1. *If there exist $\mathbb{P} \in \mathcal{S}_+^{n+p}$, and a matrix $K \in \mathbb{R}^{n\times p}$ such that*

$$\mathbb{G}^\mathsf{T} e^{\mathbb{F}^\mathsf{T} v}\mathbb{P}e^{\mathbb{F}v}\mathbb{G} - \mathbb{P} < \mathbf{0} \qquad \forall v \in [T_1^\mathcal{U}, T_2^\mathcal{U}], \tag{7.11}$$

where

$$\mathbb{F} = \begin{bmatrix} A & B \\ 0 & 0 \end{bmatrix}, \mathbb{G} = \begin{bmatrix} \mathrm{I} & 0 \\ K & 0 \end{bmatrix} \tag{7.12}$$

then Assumption 7.2 is verified.

Proof. Consider system (7.10) and set

$$f_\mathcal{K}(x_\mathcal{K}) = \begin{bmatrix} \mathbb{F}\begin{bmatrix} z \\ u \end{bmatrix} \\ -1 \end{bmatrix}, G_\mathcal{K}(x_\mathcal{K}) = \begin{bmatrix} \mathbb{G}\begin{bmatrix} z \\ u \end{bmatrix} \\ [T_1^\mathcal{U}, T_2^\mathcal{U}] \end{bmatrix}.$$

Pick the following Lyapunov function candidate for the hybrid system (7.10) defined for every $x_\mathcal{K} = (z, u, \tau_1) \in \mathbb{R}^{n+p+1}$:

$$V(x_\mathcal{K}) = \begin{bmatrix} z \\ u \end{bmatrix}^T e^{\mathbb{F}^\mathsf{T}\tau_1}\mathbb{P}e^{\mathbb{F}\tau_1}\begin{bmatrix} z \\ u \end{bmatrix}. \tag{7.13}$$

To prove the claim, we pursue a similar approach as the one in the proof of Theorem 5.1. To this end, notice that there exist two positive scalars α_1, α_2 such that

$$\alpha_1|x_\mathcal{K}|^2_{\mathcal{A}_\mathcal{K}} \le V(x_\mathcal{K}) \le \alpha_2|x_\mathcal{K}|^2_{\mathcal{A}_\mathcal{K}} \ \forall x_\mathcal{K} \in C_\mathcal{K} \cup D_\mathcal{K} \cup G_\mathcal{K}(D_\mathcal{K}) \tag{7.14}$$

where $\mathcal{A}_\mathcal{K}$ is defined in Assumption 7.2. Specially, due to the positive definiteness of $\mathbb{P}$ and the non-singularity of the matrix $e^{\mathbb{F}\tau_1}$ for every τ_1, by continuity arguments, one can set

$$\alpha_1 = \min_{\tau_1\in[0,T_2^\mathcal{U}]} \lambda_{\min}\left(e^{\mathbb{F}^\mathsf{T}\tau_1}\mathbb{P}e^{\mathbb{F}\tau_1}\right) \tag{7.15}$$

$$\alpha_2 = \max_{\tau_1 \in [0, T_2^{\mathcal{U}}]} \lambda_{\max}\left(e^{\mathbb{F}^\top \tau_1} \mathbb{P} e^{\mathbb{F}\tau_1}\right) \tag{7.16}$$

where $\lambda_{\min}(\cdot)$ and $\lambda_{\max}(\cdot)$ denote, respectively, the smallest and the largest eigenvalue of their matrix argument. By straightforward calculations one gets

$$\nabla V(x_{\mathcal{K}}) = \left(2e^{\mathbb{F}^\top \tau_1}\mathbb{P}e^{\mathbb{F}\tau_1}\begin{bmatrix} z \\ u \end{bmatrix}, [z\ u]e^{\mathbb{F}^\top \tau_1}(\mathbb{F}^\top \mathbb{P} + \mathbb{P}\mathbb{F})e^{\mathbb{F}\tau_1}\begin{bmatrix} z \\ u \end{bmatrix}\right)$$

Since the matrices $e^{\mathbb{F}\tau_1}$ and $\mathbb{F}$ commute, one has

$$\langle \nabla V(x), f_{\mathcal{K}}(x_{\mathcal{K}})\rangle = 0 \qquad \forall x_{\mathcal{K}} \in C_{\mathcal{K}}. \tag{7.17}$$

Notice that, for every $g_{\mathcal{K}} \in G_{\mathcal{K}}(x_{\mathcal{K}})$, there exists a real scalar v belonging to the interval $[T_1^{\mathcal{U}}, T_2^{\mathcal{U}}]$ such that

$$g_{\mathcal{K}} = \begin{bmatrix} \mathbb{G}\begin{bmatrix} z \\ u \end{bmatrix} \\ v \end{bmatrix}.$$

Then, for every $g_{\mathcal{K}} \in G_{\mathcal{K}}(x_{\mathcal{K}})$, one has

$$V(g_{\mathcal{K}}) - V(x_{\mathcal{K}}) = \begin{bmatrix} z \\ u \end{bmatrix}^\top \left(\mathbb{G}^\top e^{\mathbb{F}^\top v}\mathbb{P}e^{\mathbb{F}v}\mathbb{G} - e^{\mathbb{F}^\top v}\mathbb{P}e^{\mathbb{F}v}\right)\begin{bmatrix} z \\ u \end{bmatrix}$$

Moreover, whenever $x_{\mathcal{K}} \in D_{\mathcal{K}}$, from (7.10b), we have that $\tau_1 = 0$. Then, we have

$$\begin{bmatrix} z \\ u \end{bmatrix}^\top \left(\mathbb{G}^\top e^{\mathbb{F}^\top v}\mathbb{P}e^{\mathbb{F}v}\mathbb{G} - \mathbb{P}\right)\begin{bmatrix} z \\ u \end{bmatrix}$$

Hence, by virtue of relation (7.11), it follows that there exists a positive small enough scalar β such that, for every $v \in [T_1^{\mathcal{U}}, T_2^{\mathcal{U}}]$, and $\forall x_{\mathcal{K}} \in D_{\mathcal{K}}, \forall g_{\mathcal{K}} \in G_{\mathcal{K}}(x_{\mathcal{K}})$

$$V(g_{\mathcal{K}}) - V(x_{\mathcal{K}}) \leq -\beta|x_{\mathcal{K}}|^2_{\mathcal{A}_{\mathcal{K}}} \tag{7.18}$$

Now, let ϕ be a solution to (7.10). Notice that, for each $(t, j) \in \operatorname{dom}\phi$, one has $t \leq T_2^{\mathcal{U}}(j+1)$. Hence, by following the same arguments presented in the proof of Theorem 5.1, thanks to (7.17) and (7.18) it follows that the set $\mathcal{A}_{\mathcal{K}}$ is GES, hence GAS, for system (7.10). Hence, Assumption 7.2 is verified, concluding the proof. ■

7.3.3 Design Procedure

Direct computation of the gain K via Proposition 7.1 is not straightforward. In particular, from a numerical standpoint, (7.11) has two issues: it is not linear in $\mathbb{P}$ and K, and it needs to be verified for infinitely many values of v. Thus, to make the problem numerically tractable, inspired by the results presented in Chapter 5, some manipulations are needed. To this end, the following results allow to derive an LMI-based design procedure for the proposed controller.

Proposition 7.2. *If there exist a matrix $K \in \mathbb{R}^{p\times n}$, and $P_1 \in \mathcal{S}_+^n$, such that for each $v \in [T_1^{\mathcal{U}}, T_2^{\mathcal{U}}]$,*

$$\left(e^{Av} + \int_0^v e^{As} ds BK\right)^{\mathsf{T}} P_1 \left(e^{Av} + \int_0^v e^{As} ds BK\right) - P_1 < \mathbf{0} \tag{7.19}$$

then, there exists $\mathbb{P} \in \mathcal{S}_+^{n+p}$ such that the pair $(K, \mathbb{P})$ satisfies (7.11).

Proof. First of all notice that, for every real scalar v, the following identity holds

$$e^{\left(\begin{bmatrix} A & B \\ \mathbf{0} & \mathbf{0} \end{bmatrix} v\right)} = \begin{bmatrix} e^{Av} & \int_0^v e^{As} ds B \\ \mathbf{0} & \mathrm{I} \end{bmatrix} \tag{7.20}$$

and consequently

$$e^{\mathbb{F}v}\mathbb{G} = \begin{bmatrix} e^{Av} + \int_0^v e^{As} ds BK & \mathbf{0} \\ K & \mathbf{0} \end{bmatrix}. \tag{7.21}$$

Hence, since the $(2,2)$-block of the aforementioned matrix is zero, thanks to Lemma C.3, it turns out that from (7.19) there exists $\mathbb{P} \in \mathcal{S}_+^{n+p}$ such that (7.11) holds and this concludes the proof. ■

Now, we proceed to provide a condition linear in the decision variables which implies (7.19).

Proposition 7.3. *The feasibility of (7.19) follows from the feasibility of*

$$\begin{bmatrix} W + S + S^{\mathsf{T}} & -e^{Av}S - \int_0^v e^{As} ds BY \\ \bullet & -W \end{bmatrix} < \mathbf{0} \qquad v \in [T_1^{\mathcal{U}}, T_2^{\mathcal{U}}] \tag{7.22}$$

with respect to $W \in \mathcal{S}_+^n$, $S \in \mathbb{R}^{n\times n}$, and $Y \in \mathbb{R}^{p\times n}$. In particular, given any feasible solution to (7.22), $K = YS^{-1}$ and $P_1 = S^{-\mathsf{T}} W S^{-1}$ satisfy (7.19).

Proof. Notice that the feasibility of (7.19), by Projection Lemma [99], follows from the feasibility of

$$\begin{bmatrix} P_1 + \mathrm{He}(X) & -X^{\mathsf{T}} e^{Av} - X^{\mathsf{T}} \int_0^v e^{As} ds BK \\ \bullet & -P_1 \end{bmatrix} < \mathbf{0} \quad v \in [T_1^{\mathcal{U}}, T_2^{\mathcal{U}}] \tag{7.23}$$

where $X \in \mathbb{R}^{n\times n}$. Now, by setting $X^{-1} = S$, $S^{\mathsf{T}} P_1 S = W$, and $KS = Y$ and by pre-and-post multiplying the left-hand side of (7.23) by $\mathrm{diag}(S^{\mathsf{T}}, S^{\mathsf{T}})$ and $\mathrm{diag}(S, S)$ provides the left-hand side of (7.22) and since the above transformations are invertible, the result is established. ■

Proposition 7.3 provides a sufficient condition to (7.19), which is linear in the decision variables W, Y, and S. In particular, the above result ensures that given (W, Y, S) such that

(7.22) holds, then $K = YS^{-1}$ lets Assumption (7.2) hold.

Nevertheless, (7.22) still has to be verified for infinitely many values of v. To overcome such a drawback, we proceed in a similar way as in Chapter 5. Namely, by building a suitable polytopic embedding, we derive a finite number of conditions whose satisfaction yields the satisfaction of (7.23). To this end, consider the following preliminary result, whose proof is given in Appendix C.

Lemma 7.2. Let v be a real scalar belonging to a given compact interval $\mathcal{I}$, and let Ω_1 and Ω_2 be two real constant matrices. Let

$$X_1 = \begin{bmatrix} R_1 & Q_1 \\ U_1 & L_1 \end{bmatrix}, X_2 = \begin{bmatrix} R_2 & Q_2 \\ U_2 & L_2 \end{bmatrix}, \dots, X_\nu = \begin{bmatrix} R_\nu & Q_\nu \\ U_\nu & L_\nu \end{bmatrix}$$

be matrices such that for each $v \in \mathcal{I}$,

$$\exp\left(\begin{bmatrix} \Omega_1 & \Omega_2 \\ \mathbf{0} & \mathbf{0} \end{bmatrix} v\right) \in \operatorname{co}\left\{X_1, X_2, \dots, X_\nu\right\}. \tag{7.24}$$

Then, for each $v \in \mathcal{I}$, the following identities hold:

$$\begin{bmatrix} e^{\Omega_1 v} & \int_0^v e^{\Omega_1 s} ds \Omega_2 \end{bmatrix} \in \operatorname{co}\left\{\begin{bmatrix} R_1 & Q_1 \end{bmatrix}, \begin{bmatrix} R_2 & Q_2 \end{bmatrix}, \dots, \begin{bmatrix} R_\nu & Q_\nu \end{bmatrix}\right\} \tag{7.25}$$

□

The above result is rather general, since it is not based on a specific polytopic embedding of the exponential matrix in (7.24). Thus, to achieve the desired task, any of the technique proposed in the literature can be adopted. In this dissertation, we rely on technique exposed in Chapter 5.

Now we are in position to state the following design result.

Proposition 7.4. *Let*

$$X_1 = \begin{bmatrix} R_1 & Q_1 \\ U_1 & L_1 \end{bmatrix}, X_2 = \begin{bmatrix} R_2 & Q_2 \\ U_2 & L_2 \end{bmatrix}, \dots, X_\nu = \begin{bmatrix} R_\nu & Q_\nu \\ U_\nu & L_\nu \end{bmatrix} \tag{7.26}$$

be matrices such that for each $v \in [T_1^{\mathcal{U}}, T_2^{\mathcal{U}}]$,

$$\exp\left(\begin{bmatrix} A & B \\ \mathbf{0} & \mathbf{0} \end{bmatrix} v\right) \in \operatorname{co}\left\{X_1, X_2, \dots, X_\nu\right\}. \tag{7.27}$$

If there exist $W \in \mathcal{S}_+^n$, $S \in \mathbb{R}^{n\times n}$, *and* $Y \in \mathbb{R}^{p\times n}$ *such that for every* $i = 1, \dots, \nu$

$$\begin{bmatrix} W + S + S^\mathsf{T} & -R_i S - Q_i Y \\ \bullet & -W \end{bmatrix} < \mathbf{0}, \tag{7.28}$$

then $K = YS^{-1}$ *ensures Assumption 7.1.*

Proof. First of all, according to Lemma 7.2, there exist positive functions $\lambda_1(v), \ldots, \lambda_\nu(v)$, such that for each $v \in [T_1^{\mathcal{U}}, T_2^{\mathcal{U}}]$

$$e^{Av} = \sum_{i=1}^{\nu} \lambda_i(v) R_i, \int_0^v e^{As} ds B = \sum_{i=1}^{\nu} \lambda_i(v) Q_i \tag{7.29}$$

with $\sum_{i=1}^{\nu} \lambda_i(v) = 1$. Thus, the left-hand side of (7.22) turns in

$$\sum_{i=1}^{\nu} \lambda_i(v) \begin{bmatrix} W + S + S^{\mathsf{T}} & -R_i S - Q_i Y \\ \bullet & -W \end{bmatrix}. \tag{7.30}$$

Hence by the virtue of Proposition 7.3 the result is proven. ■

7.4 Numerical example

Example 7.1. Consider the linearized model for the unstable batch reactor in [58], which is described by the following data:

$$
\begin{aligned}
A &= \begin{bmatrix} 1.38 & -0.208 & 6.71 & -5.68 \\ -0.581 & -4.29 & 0 & 0.675 \\ 1.07 & 4.27 & -6.65 & 5.89 \\ 0.048 & 4.27 & 1.34 & -2.1 \end{bmatrix}, B = \begin{bmatrix} 0 & 0 \\ 5.68 & 0 \\ 1.14 & -3.15 \\ 1.14 & 0 \end{bmatrix} \\
M &= \begin{bmatrix} 1 & 0 & 1 & -1 \\ 0 & 1 & 0 & 0 \end{bmatrix}
\end{aligned}
\tag{7.31}
$$

and assume $T_1^{\mathcal{O}} = T_1^{\mathcal{U}} = T_1 = 0.1$ and $T_2^{\mathcal{O}} = T_2^{\mathcal{U}} = T_2 = 0.9$. As a first step, by relying on the apparatus illustrated in Section 5.4, we design the observer gain L to let Assumption 7.1 hold. In particular, we obtain

$$
L = \begin{bmatrix} 0.8618 & -0.1012 \\ 0.0001516 & 1 \\ 0.131 & 0.277 \\ -0.006379 & 0.1765 \end{bmatrix}
$$

Then, as a second step, to let Assumption 7.2 hold we design the controller gain K via Proposition 7.4. In particular, by building on the polytopic embedding proposed in Section 5.4.1, one gets

$$
K = \begin{bmatrix} 0.19355 & -0.17442 & 0.094692 & -0.23368 \\ 1.2263 & 0.087818 & 0.85837 & -0.53913 \end{bmatrix}.
$$

Figure 7.2 shows the evolution of the plant state z and Figure 7.3 shows the evolution of the observer state $\hat{z}$ projected onto ordinary time. While Figure 7.4 and Figure 7.5 show, respectively, the evolution of the input u feeding the plant and the evolution two timers τ_1 and τ_2, still projected onto ordinary time. In this simulation, $z(0,0) = (1,1,1,1), \hat{z}(0,0) = (0,0,0,0), u(0,0) = (0,0), \hat{u}(0,0) = (0,0), \tau_1(0,0) = \tau_2(0,0) = T_2$, and the sampling instants are chosen randomly according to a uniform distribution. Simulations show the effectiveness of the proposed approach, by stressing that the stabilization is achieved despite the lack of synchronism between the output sampling and input updating, as Figure 7.5 suggests. Moreover, it is interesting to notice that, according to the initialization operated for τ_1 and τ_2, the control system runs in open-loop for the first T_2 units of time, as underlined by Figure 7.3 and Figure 7.4.

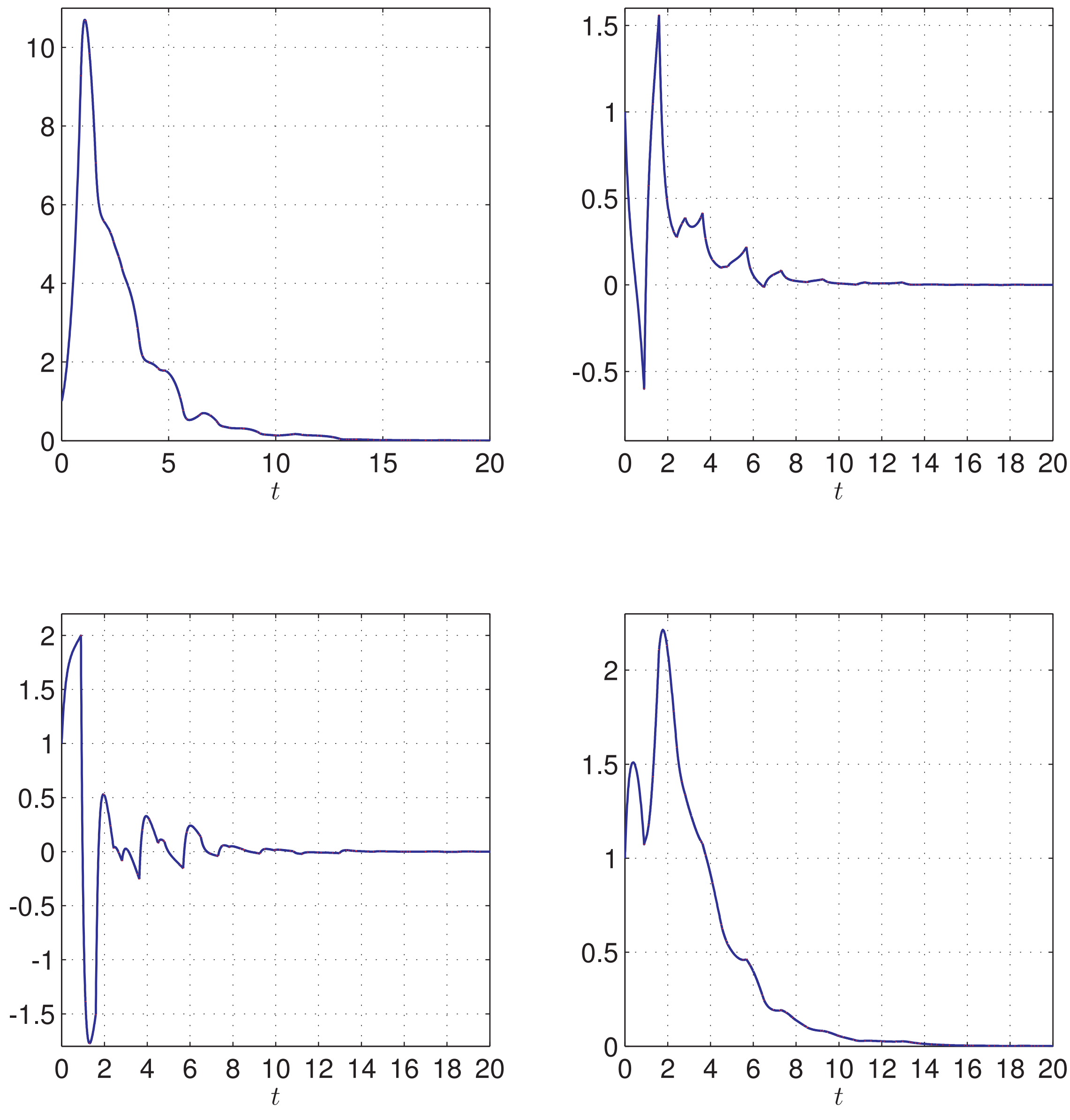


Figure 7.2: The evolution of the plant state z.

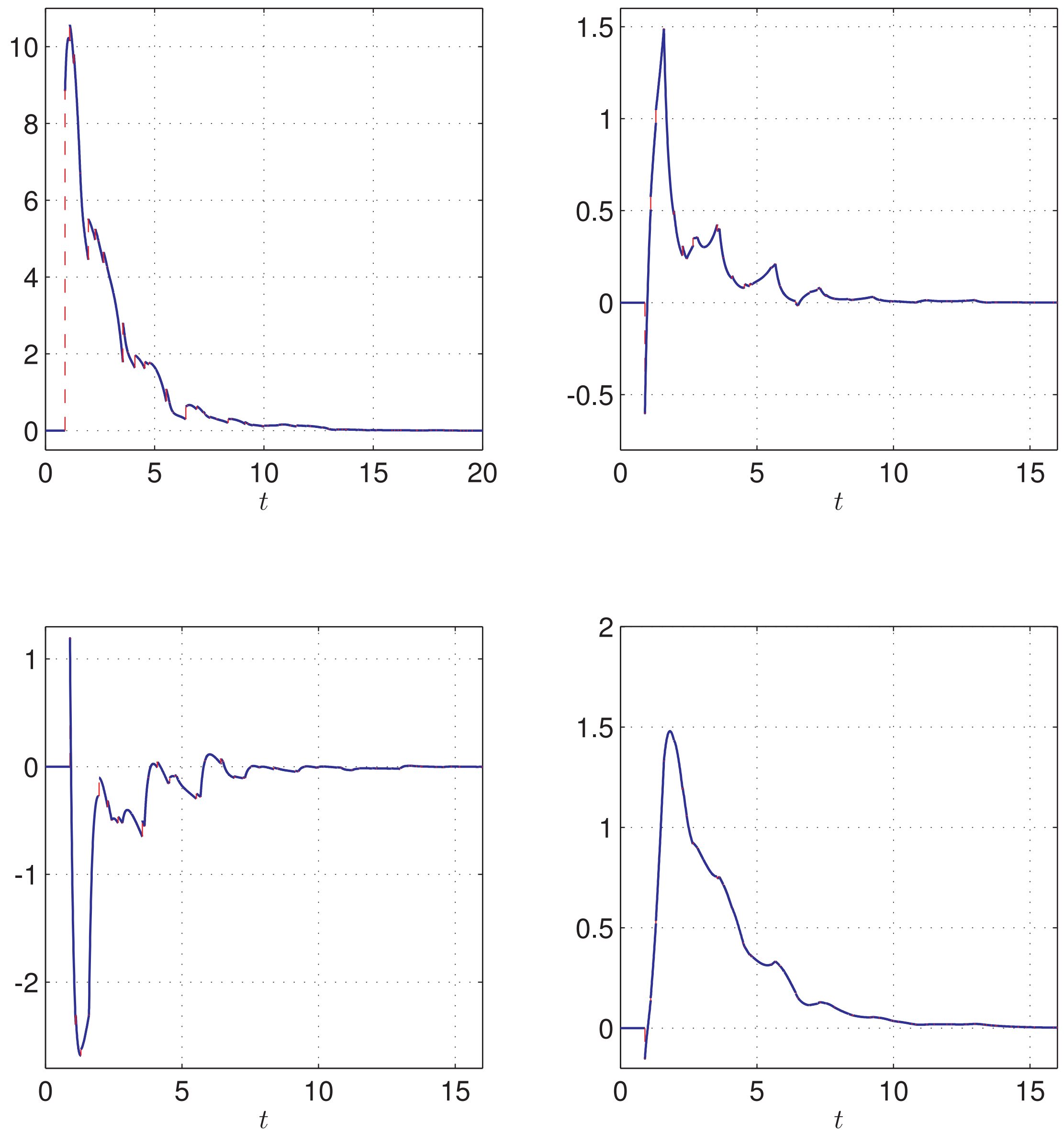


Figure 7.3: The evolution of observer state $\hat{z}$ projected onto ordinary time.

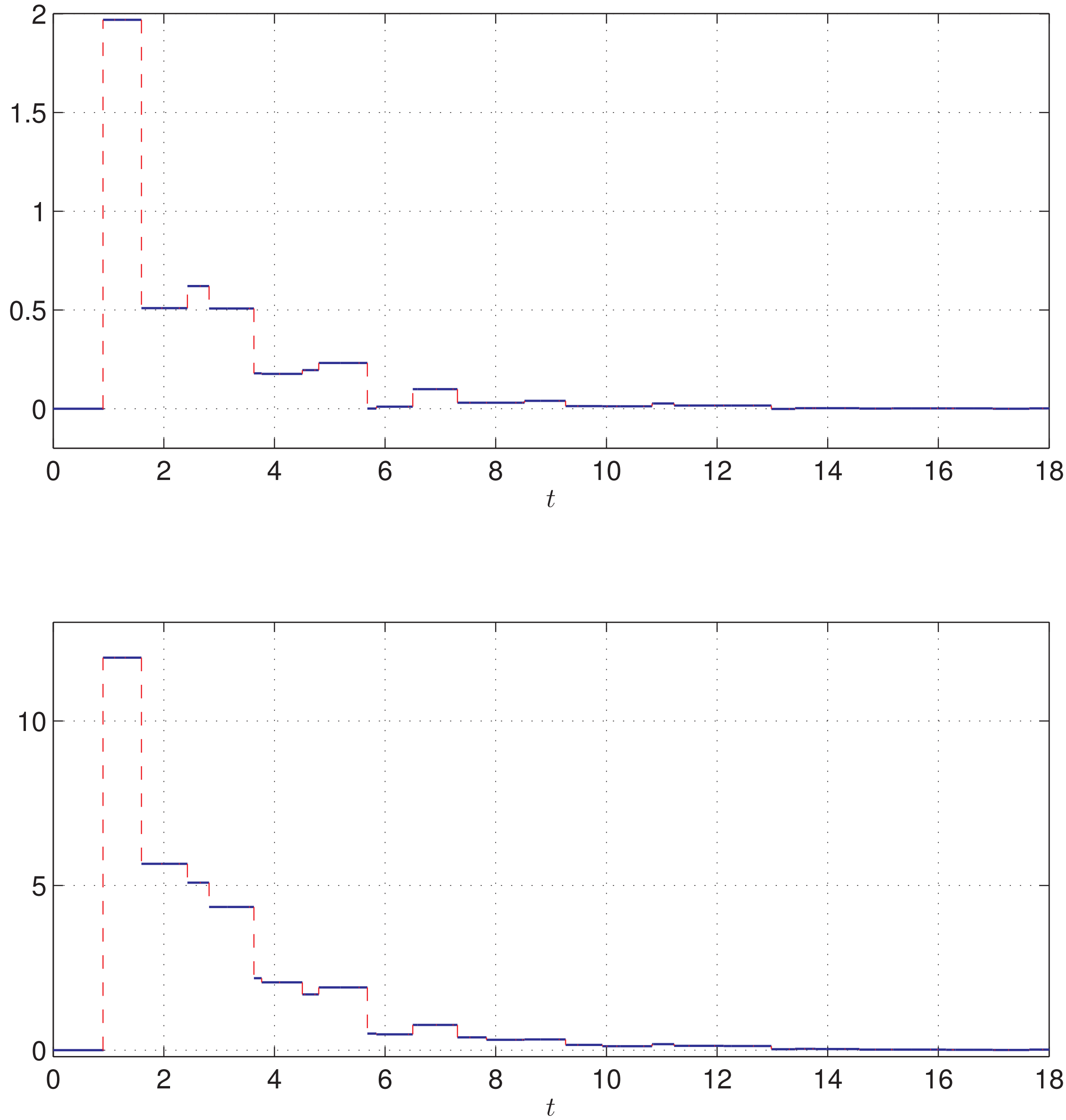


Figure 7.4: The evolution of u projected onto ordinary time.

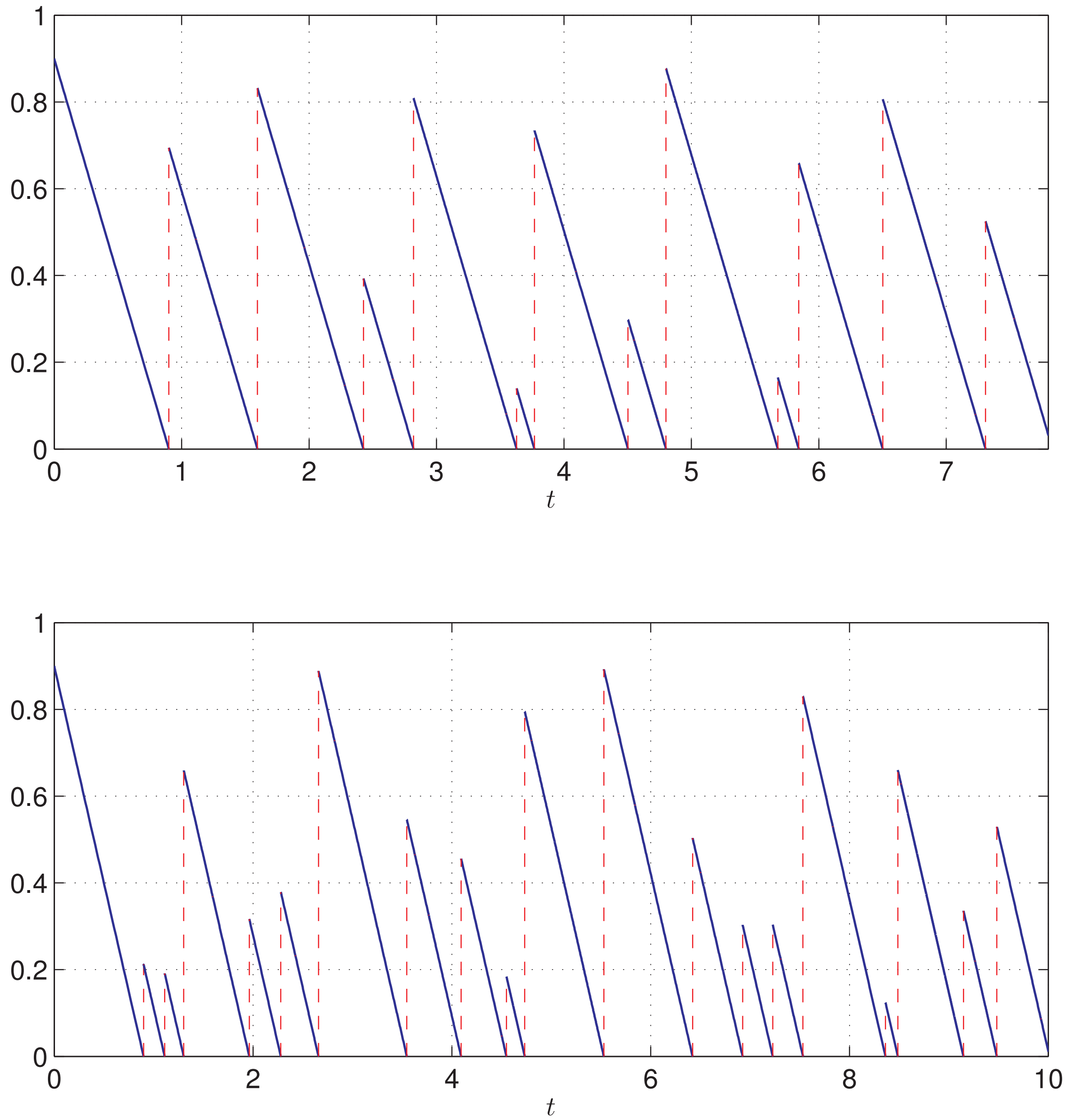


Figure 7.5: The evolution of τ_1 (above) and τ_2 (below) projected onto ordinary time.

7.5 Comments and Conclusion

In this chapter, we shown how the measurement-triggered observer proposed in Chapter 5 can be used to asymptotically stabilize a linear plant in the presence of sporadic measurements even when the plant input is not accessible at any time, provided that the controller is aware whenever a new sample is sent to the plant. The proposed approach builds on a separation principle, which due to the homogeneity of the resulting hybrid system leads to a global result. Moreover, a numerically tractable design, based on the solution to certain LMIs was provided. Finally, the effectiveness of the described methodology is shown in an example.

One of the main advantages of the proposed approach consists in avoiding the need of seeking for a Lyapunov function for the whole closed-loop system to certify asymptotic stability, which is a nontrivial problem. This is a worthwhile feature, which is enabled by the use of the general and powerful framework proposed in [56] to study hybrid dynamical systems. In particular, building on [56] allows to mimic the standard arguments adopted to establish stability properties for upper triangular continuous-time or discrete-time nonlinear systems. The pursued approach also brings outs that, in the considered setting, using an observer-based controller enables to achieve closed-loop asymptotic stability without assuming any correlation between the output sampling events and the control input updating events.

One should be aware that the same approach could be considered to stabilize nonlinear plants, as long as one is able to build, for the considered case, an observer to reconstruct the plant state in the presence of sporadic measurement and a state feedback controller to stabilize the plant in the presence of sporadic input access. However, in this setting, the considered separation principle could allow to establish only local results, as in the more general case considered in [122], unless the plant to stabilize gives rise to an homogeneous closed-loop system. Whenever homogeneity does not hold, as suggested in the proof of Theorem 7.1, an estimate of the basin of attraction of the closed-loop system can be determined by seeking for a set from which the initialization of the closed-loop system leads to bounded solutions. This is certainly a difficult problem in general. In addition, often estimates of basin of attraction are built via the construction of a Lyapunov function for the closed-loop system. On the other hand, the knowledge of a (strict) Lyapunov function enables itself to conclude on asymptotic stability of the closed-loop system, making the use of a separation principle worthless, unless only a weak Lyapunov function is available.

As in Chapter 5, an interesting direction of research consists in the search of other Lyapunov functions to ensure the satisfaction of Assumption 7.2, with the aim of reducing the conservatism and the computational burden of the proposed polytopic embedding-based design procedure.

Concerning comparisons between the approach we illustrated and the other approaches usually considered in the literature to deal with stability and stabilization in the presence of asynchronous sampling, it is worthwhile to observe that adopting a discrete-time approach, as the one often considered in the literature of sampled-data and networked control systems;

see, *e.g.*, [29] and the references there in, does not seem suitable in this setting due to the asynchronicity of the output sampling events and the control input updating events. Indeed, such an approach rests on the construction of a discrete-time model of the closed-loop system, process that does not appear doable in our setting due to multiple asynchronous jumps.

CONCLUSION OF PART II

Concluding Remarks

In this part of this thesis, we provided two observer schemes to exponentially estimate the state of a continuous-time LTI system in the presence of sporadic measurements. In addition, building on the first considered observer scheme, an observer-based controller scheme is proposed to asymptotically stabilize a continuous-time LTI system in the presence of both sporadic measurements and input access. For such a scheme, a separation principle was shown. The pursued approach hinges upon the hybrid system framework in [56] and leads to computationally tractable conditions for the design of the resulting observation/controller schemes.

Perspectives and Future Outlook

The work presented within this part has the role to prepare the stage for several interesting extensions. In this sense, the results included in this part do not represent an ended work. In particular, as previously mentioned, the extension of the observer in Chapter 6 to multi-output linear plants with asynchronous channels in the presence of sporadic measurements is currently under investigation. A likewise interesting extension consists of the construction of an observer-based controller, as in Chapter 7, to account event for asynchronous input channels in multi-inputs plants. This latter lines of research suggests that a possible extension of the work presented in this dissertation concerns the construction of an observer-based controller in the case of networked control systems; see [129]. Indeed, in this setting actuators and sensors are grouped in different nodes that, respectively, grant their access and transmit data sporadically. Thus, such a situation can be addressed, with some extra work, merging the ideas in Chapter 7 and the observer presented in Chapter 6, though adapted for the case of asynchronous multi-outputs sampling. However, observe that to rest on the controller architecture proposed in Chapter 7, as assumed therein, one would need to ensure that whenever a new control sample is sent to the plant, the controller instantaneously updates its internal variables to keep track on the plant input, so as to build a correct estimation of the plant state. This assumption should not entail a severe constraint in practical implementations and usually considered in the literature of networked control systems; see, *e.g.*, [62]. For instance, a packet acknowledgment mechanism, as the one implemented in the TCP protocol, would enable to effectively ensure such an assumption; see [113].

In the framework of networked control systems, an aspect that deserves investigations pertains to the presence of time-delays in the considered input and output channels, which is a well acknowledged in the literature of networked control systems; see [62]. Also such a problematic could be addressed in a hybrid systems setting via the notion of hybrid system with memory illustrated in [86], although the extension does not appear straightforward.

Concerning genuine observer design in the presence of sporadic measurements, an interesting aspect consists of coupling some performance requirements to the synthesis procedures proposed within this dissertation. For instance, as already done, *e.g.*, in [45], one can envision the derivation of design strategies guaranteeing a given exponential decay-rate for the estimation error, and/or ensuring some performance in terms of attenuation of exogenous signals.

Appendices

APPENDIX B

B.1 Extreme matrices of Example 5.2

$$X_1 = \begin{bmatrix} 0.12242 & 0.14812 & 0.39226 & -0.011693 \\ 0.31962 & 0.09904 & -0.023385 & 0.22076 \\ -0.8079 & 0.41564 & -0.26983 & 0.1715 \\ 0.48828 & -0.4649 & 0.34301 & -0.34247 \end{bmatrix}$$

$$X_2 = \begin{bmatrix} 0.12242 & -0.70693 & -0.46279 & -0.86674 \\ 0.31962 & -1.6111 & -1.7335 & -1.4893 \\ -0.8079 & 1.2707 & 0.58522 & 1.0266 \\ 0.48828 & 1.2452 & 2.0531 & 1.3676 \end{bmatrix}$$

$$X_3 = \begin{bmatrix} 0.45744 & 1.2652 & 1.32 & 0.80905 \\ 0.91237 & 2.0755 & 1.6181 & 1.6729 \\ -1.0219 & -0.29808 & -0.86258 & -0.35287 \\ 0.10958 & -1.7277 & -0.70574 & -1.2702 \end{bmatrix}$$

$$X_4 = \begin{bmatrix} 0.45744 & 0.41019 & 0.46498 & -0.045998 \\ 0.91237 & 0.36545 & -0.091995 & -0.037207 \\ -1.0219 & 0.55697 & -0.0075306 & 0.50218 \\ 0.10958 & -0.017582 & 1.0044 & 0.43986 \end{bmatrix}$$

$$X_5 = \begin{bmatrix} -0.16909 & 0.31446 & 0.11447 & 0.14779 \\ 0.33335 & 0.12649 & 0.29558 & 0.095588 \\ 0.066639 & -0.18111 & -0.28356 & 0.018883 \\ -0.39999 & 0.1044 & 0.037766 & -0.064684 \end{bmatrix}$$

$$X_6 = \begin{bmatrix} -0.16909 & -0.54059 & -0.74058 & -0.70726 \\ 0.33335 & -1.5836 & -1.4145 & -1.6145 \\ 0.066639 & 0.67394 & 0.57149 & 0.87393 \\ -0.39999 & 1.8145 & 1.7479 & 1.6454 \end{bmatrix}$$

$$X_7 = \begin{bmatrix} 0.16593 & 1.4316 & 1.0422 & 0.96854 \\ 0.92609 & 2.103 & 1.9371 & 1.5477 \\ -0.1474 & -0.89483 & -0.87631 & -0.50549 \\ -0.77869 & -1.1584 & -1.011 & -0.99245 \end{bmatrix}$$

$$X_8 = \begin{bmatrix} 0.16593 & 0.57653 & 0.18719 & 0.11349 \\ 0.92609 & 0.3929 & 0.22697 & -0.16237 \\ -0.1474 & -0.039783 & -0.021257 & 0.34956 \\ -0.77869 & 0.55172 & 0.69912 & 0.71765 \end{bmatrix}$$

$$X_9 = \begin{bmatrix} 0.8232 & -0.35304 & 0.1821 & -0.057382 \\ -0.59132 & 0.70844 & -0.11476 & 0.42038 \\ -0.47896 & 0.29686 & 0.6411 & -0.23828 \\ 1.0703 & -0.95551 & -0.47655 & -0.13231 \end{bmatrix}$$

$$X_{10} = \begin{bmatrix} 0.8232 & -1.2081 & -0.67295 & -0.91243 \\ -0.59132 & -1.0017 & -1.8249 & -1.2897 \\ -0.47896 & 1.1519 & 1.4962 & 0.61677 \\ 1.0703 & 0.75459 & 1.2335 & 1.5778 \end{bmatrix}$$

$$X_{11} = \begin{bmatrix} 1.1582 & 0.76408 & 1.1099 & 0.76336 \\ 0.0014298 & 2.6849 & 1.5267 & 1.8725 \\ -0.693 & -0.41686 & 0.048357 & -0.76265 \\ 0.69157 & -2.2183 & -1.5253 & -1.0601 \end{bmatrix}$$

$$X_{12} = \begin{bmatrix} 1.1582 & -0.090972 & 0.25481 & -0.091687 \\ 0.0014298 & 0.97485 & -0.18337 & 0.16241 \\ -0.693 & 0.43819 & 0.90341 & 0.092402 \\ 0.69157 & -0.5082 & 0.1848 & 0.65002 \end{bmatrix}$$

$$X_{13} = \begin{bmatrix} 0.53169 & -0.18669 & -0.095689 & 0.1021 \\ -0.57759 & 0.73589 & 0.2042 & 0.29521 \\ 0.39558 & -0.29989 & 0.62738 & -0.3909 \\ 0.18201 & -0.38621 & -0.78179 & 0.14548 \end{bmatrix}$$

$$X_{14} = \begin{bmatrix} 0.53169 & -1.0417 & -0.95074 & -0.75295 \\ -0.57759 & -0.97421 & -1.5059 & -1.4149 \\ 0.39558 & 0.55516 & 1.4824 & 0.46415 \\ 0.18201 & 1.3239 & 0.92831 & 1.8556 \end{bmatrix}$$

$$X_{15} = \begin{bmatrix} 0.86671 & 0.93042 & 0.83208 & 0.92285 \\ 0.015156 & 2.7124 & 1.8457 & 1.7473 \\ 0.18154 & -1.0136 & 0.034631 & -0.91527 \\ -0.19669 & -1.649 & -1.8305 & -0.78229 \end{bmatrix}$$

$$X_{16} = \begin{bmatrix} 0.86671 & 0.075374 & -0.022973 & 0.067796 \\ 0.015156 & 1.0023 & 0.13559 & 0.037246 \\ 0.18154 & -0.15857 & 0.88968 & -0.060218 \\ -0.19669 & 0.061102 & -0.12044 & 0.92781 \end{bmatrix}$$

This page is intentionally left blank.

APPENDIX C

Proof of Claim 7.1. Pick any maximal solution $\phi = (\phi_1, \phi_2)$ to system $\mathcal{H}_{\mathcal{A}_{\tilde{u}}}$, where

$$(t,j) \mapsto \phi_1(t,j) = (z(t,j), u(t,j), \tau_1(t,j))$$
$$(t,j) \mapsto \phi_2(t,j) = (\varepsilon(t,j), 0, \tau_2(t,j)).$$

According to the properties of the domain of the solutions to (7.7) shown in Section 7.2.2, it is straightforward to show that there exists a solution φ to system (7.9), with $\sup_t \operatorname{dom}\varphi = \sup_t \operatorname{dom}\phi$, and such that for every $(t,j) \in \operatorname{dom}\varphi$ there exists $s \in \mathbb{N}, s \geq j\colon (t,s) \in \operatorname{dom}\phi$ and $\varphi(t,j) = [\varepsilon(t,s), \tau_2(t,s)]$. Loosely speaking, φ flows whenever ϕ flows and only jumps whenever ϕ jumps due to $\phi(t,s) \in D_{2e}$. Moreover, notice that for every $(z, u, \tau_1, \varepsilon, \tilde{u}, \tau_2) \in (C_e \cap \mathcal{A}_{\tilde{u}}) \cup (D_e \cap \mathcal{A}_{\tilde{u}}) \cup G_e(D_e \cap \mathcal{A}_{\tilde{u}})$, one has $|(z, u, \tau_1, \varepsilon, \tilde{u}, \tau_2)|_{\mathcal{A}_\varepsilon} = |(\varepsilon, \tau_2)|_{\mathcal{A}_o}$. Now, from stability of $\mathcal{A}_o$ for system (7.9), one has for every $(t,j) \in \operatorname{dom}\varphi$, $|\varphi(t,j)|_{\mathcal{A}_o} \leq |\varphi(0,0)|_{\mathcal{A}_o}$, and by construction for all $(t,s) \in \operatorname{dom}\phi$ there exists $j \in \mathbb{N}$ with $j \leq s$, such that $(t,j) \in \operatorname{dom}\varphi$ and

$$|\phi(t,s)|_{\mathcal{A}_\varepsilon} = |\phi_2(t,s)|_{\mathcal{A}_o} = |\varphi(t,j)|_{\mathcal{A}_o}.$$

Then, stability of $\mathcal{A}_\varepsilon$ is proven. Concerning global attractivity, pick any maximal solution ϕ to $\mathcal{H}_{\mathcal{A}_{\tilde{u}}}$, which is complete, and suppose that the set $\mathcal{A}_\varepsilon$ is not attractive for ϕ. Then, there exists $h > 0$ such that for every positive scalar T, $t + s \geq T$ and $(t,s) \in \operatorname{dom}\phi$ implies $|\phi(t,s)|_{\mathcal{A}_\varepsilon} \geq h$. Now, pick $(t,j) \in \operatorname{dom}\varphi$ and such that $t + j \geq T$. Then, there exists $s^* \geq j$ such that $|\varphi(t,j)|_{\mathcal{A}_o} = |\phi(t,s^*)|_{\mathcal{A}_\varepsilon}$, but since $t + s^* \geq t + j \geq T$ one has $|\varphi(t,j)|_{\mathcal{A}_o} \geq h$, which contradicts the fact that $\mathcal{A}_o$ is globally attractive for the hybrid system (7.9), and this concludes the proof. ■

Lemma C.1. Given a hybrid system $\mathcal{H} = (C, F, D, G)$ and a compact set J. Let $\mathcal{H}_J$ be the restriction of $\mathcal{H}$ to J. Assume that, for some set $S \subset \operatorname{Int}J$, each $\phi \in \mathcal{S}_{\mathcal{H}_J}(S)$ is such that $\operatorname{rge}\phi \subset \operatorname{Int}J$. Then, $\mathcal{S}_{\mathcal{H}}(S) \subset \mathcal{S}_{\mathcal{H}_{\mathcal{J}}}(S)$.

Proof. To establish the result, it suffices to show that each $\psi \in \mathcal{S}_{\mathcal{H}}(S)$ is a solution to $\mathcal{H}_J$.

Indeed, since each solution to $\mathcal{H}_J$ is a solution to $\mathcal{H}$, maximality of such solutions for $\mathcal{H}_J$ directly follows from the fact that they are maximal for $\mathcal{H}$.

By contradiction, let us assume that there exists $\psi \in \mathcal{S}_{\mathcal{H}}(S)$, which is not a solution to $\mathcal{H}_J$. Then, by definition of solution, since $\mathcal{H}$ and $\mathcal{H}_J$ have the same dynamics, and $\psi(0,0) \in S \cap (C \cup D) \subset \mathrm{Int}J$ the only possibility for ψ not being a solution to $\mathcal{H}_J$ is that ψ eventually leaves J. Let us assume that ψ leaves J via a jump. Then, there exist $(t,j) \in \mathrm{dom}\,\psi$ such that $(t, j+1) \in \mathrm{dom}\,\psi$, $\psi(t,j) \in \mathrm{Int}J$, $\psi(t,j+1) \in G(\psi(t,j)) \notin J$. Thus, this implies that there exists a solution ϕ to $\mathcal{H}_J$ starting in $\mathrm{Int}J$ that leaves J, and this is not possible by assumption.

Let us assume that ψ leaves J by flowing. Then, there exists $(s,j) \in \mathrm{dom}\,\psi$ such that $\psi(s,j) \notin J$. By continuity of the function $t \mapsto \psi(t,j)$ over $[t_j, t_{j+1}]$, there exists $s^\star \in [t_j, t_{j+1}]$ such that $\psi(s^\star, j) \notin \mathrm{Int}J$. This implies that there exists a solution to $\mathcal{H}_J$ that leaves $\mathrm{Int}J$, and this contradicts the hypothesis. Then each $\psi \in \mathcal{S}_{\mathcal{H}}(S)$ is a solution to $\mathcal{H}_J$, concluding the proof. ∎

Lemma C.2. Let ϕ be a solution to (7.7). For each $\lambda > 0$, let $M_\lambda = \mathrm{diag}(\lambda \mathrm{I}, \lambda \mathrm{I}, 1, \lambda \mathrm{I}, \lambda \mathrm{I}, 1)$. For each $(t,j) \in \mathrm{dom}\,\phi$ consider the function $\psi(t,j) = M_\lambda \phi(t,j)$. Then, ψ is a solution to (7.7).

Proof. The proof follows the lines of [56, Lemma 9.3.]. In particular, we show that the hybrid system (7.7) is homogeneous of degree zero with respect to the standard nonproper dilation M_λ defined above. For each $\lambda > 0$ and for each $x_e \in C_e \cup D_e$, one gets

$$F_e(M_\lambda x_e) = M_\lambda F_e(x_e), G_e(M_\lambda x_e) = M_\lambda G_e(x_e)$$

moreover, $M_\lambda C_e = C_e, M_\lambda D_e = D_e$. Now, pick any solution ϕ to (7.7), and notice that obviously $M_\lambda \phi$ is a hybrid arc, in particular one has $\mathrm{dom}\,\phi = \mathrm{dom}\, M_\lambda \phi$. To conclude, pick $(t,j) \in \mathrm{dom}\,\phi$. Hence, if $\phi(t,j) \in C_e$ then $M_\lambda \phi(t,j) \in C_e$, while if $\dot{\phi}(t,j) = F_e(\phi(t,j))$ then $M_\lambda \dot{\phi}(t,j) = F_e(M_\lambda \phi(t,j))$. Furtheremore, If $\phi(t,j) \in D_e$ then $M_\lambda \phi(t,j) \in D_e$, and if $\phi(t,j+1) \in G_e(\phi(t,j))$ then $M_\lambda \phi(t,j+1) \in G_e(M_\lambda \phi(t,j))$. Thus $M_\lambda \phi$ is a solution to (7.7). ∎

Proof of Lemma 7.2. First notice that, according to (7.20), it follows

$$e^{\begin{bmatrix}\Omega_1 & \Omega_2\\ \mathbf{0} & \mathbf{0}\end{bmatrix}v} = \sum_{i=1}^{\nu} \lambda_i(v) X_i = \begin{bmatrix} e^{\Omega_1 v} & \int_0^v e^{\Omega_1 s} ds \Omega_2 \\ \mathbf{0} & \mathrm{I} \end{bmatrix} \tag{C.1}$$

where X_i are some suitable matrices. Thus, by partitioning every X_i as follows

$$X_i = \begin{bmatrix} R_i & Q_i \\ U_i & L_i \end{bmatrix}$$

one has

$$\begin{bmatrix} e^{\Omega_1 v} & \int_0^v e^{\Omega_1 s} ds \Omega_2 \\ \mathbf{0} & \mathrm{I} \end{bmatrix} = \begin{bmatrix} \sum_{i=1}^{\nu} \lambda_i(v) R_i & \sum_{i=1}^{\nu} \lambda_i(v) Q_i \\ \sum_{i=1}^{\nu} \lambda_i(v) U_i & \sum_{i=1}^{\nu} \lambda_i(v) L_i \end{bmatrix} \tag{C.2}$$

and this finishes the proof. ■

Lemma C.3. Let v be a real scalar belonging to a given compact interval $\mathcal{I}$. Let $\Omega_1(v) \in \mathbb{R}^{n_1 \times n_1}$, $\Omega_2(v) \in \mathbb{R}^{n_1 \times n_2}$, and $\Omega_3(v) \in \mathbb{R}^{n_2 \times n_2}$ be given real matrices of suitable dimensions whose entries depend continuously on v. If there exist $P_1 \in \mathcal{S}_+^{n_1}, P_2 \in \mathcal{S}_+^{n_2}$ such that

$$\Omega_1^{\mathsf{T}}(v) P_1 \Omega_1(v) - P_1 < \mathbf{0} \tag{C.3}$$

$$\Omega_3^{\mathsf{T}}(v) P_2 \Omega_3(v) - P_2 < \mathbf{0} \tag{C.4}$$

then there exist two constant symmetric positive definite matrices F_1, F_2 such that for every $v \in \mathcal{I}$

$$\begin{bmatrix} \Omega_1(v) & \Omega_2(v) \\ \mathbf{0} & \Omega_3(v) \end{bmatrix}^T \begin{bmatrix} F_1 & \mathbf{0} \\ \mathbf{0} & F_2 \end{bmatrix} \begin{bmatrix} \Omega_1(v) & \Omega_2(v) \\ \mathbf{0} & \Omega_3(v) \end{bmatrix} - \begin{bmatrix} F_1 & \mathbf{0} \\ \mathbf{0} & F_2 \end{bmatrix} < \mathbf{0}.$$

Proof. First of all denote

$$\begin{aligned} -Q_1(v) &= \Omega_1^{\mathsf{T}}(v) P_1 \Omega_1(v) - P_1 \\ -Q_2(v) &= \Omega_3^{\mathsf{T}}(v) P_2 \Omega_3(v) - P_2. \end{aligned}$$

Now, let γ a positive scalar to be selected later and consider the following expression

$$\begin{aligned} &\begin{bmatrix} \Omega_1(v) & \Omega_2(v) \\ \mathbf{0} & \Omega_3(v) \end{bmatrix}^T \begin{bmatrix} P_1 & \mathbf{0} \\ \mathbf{0} & \gamma P_2 \end{bmatrix} \begin{bmatrix} \Omega_1(v) & \Omega_2(v) \\ \mathbf{0} & \Omega_3(v) \end{bmatrix} - \begin{bmatrix} P_1 & \mathbf{0} \\ \mathbf{0} & \gamma P_2 \end{bmatrix} = \\ &\begin{bmatrix} -Q_1(v) & \Omega_1^{\mathsf{T}}(v) P_1 \Omega_2(v) \\ \bullet & \Omega_2^{\mathsf{T}}(v) P_1 \Omega_2(v) - \gamma Q_2(v) \end{bmatrix}. \end{aligned} \tag{C.5}$$

Then by Schur complement the above right-hand side matrix is negative definite if and only if

$$\Omega_2^{\mathsf{T}}(v) P_1 \Omega_2(v) - \gamma Q_2(v) + (\Omega_1^{\mathsf{T}}(v) P_1 \Omega_2(v))^{\mathsf{T}} Q_1(v)^{-1} (\Omega_1^{\mathsf{T}}(v) P_1 \Omega_2(v)) < \mathbf{0} \tag{C.6}$$

that is

$$\gamma Q_2(v) > \Omega_2^{\mathsf{T}}(v) \left(P_1 + P_1 \Omega_1(v) Q_1(v)^{-1} \Omega_1^{\mathsf{T}}(v) P_1 \right) \Omega_2(v).$$

which is equivalent to

$$\gamma \mathrm{I} > Q_2^{-\frac{1}{2}}(v) \Omega_2^{\mathsf{T}}(v) \left(P_1 + P_1 \Omega_1(v) Q_1(v)^{-1} \Omega_1^{\mathsf{T}}(v) P_1 \right) \Omega_2(v) Q_2^{-\frac{1}{2}}(v).$$

The latter is satisfied if for each $v \in \mathcal{I}$

$$\gamma > \lambda_{\max} \left(Q_2^{-\frac{1}{2}}(v) \Omega_2^{\mathsf{T}}(v) \left(P_1 + P_1 \Omega_1(v) Q_1(v)^{-1} \Omega_1^{\mathsf{T}}(v) P_1 \right) \Omega_2(v) Q_2^{-\frac{1}{2}}(v) \right).$$

Thus, by continuity arguments, picking

$$\gamma > \max_{v \in \mathcal{I}} \lambda_{\max} \left(Q_2^{-\frac{1}{2}}(v) \Omega_2^{\mathsf{T}}(v) \left(P_1 + P_1 \Omega_1(v) Q_1(v)^{-1} \Omega_1^{\mathsf{T}}(v) P_1 \right) \Omega_2(v) Q_2^{-\frac{1}{2}}(v) \right)$$

brings the desired result, with $F_1 = P_1$ and $F_2 = \gamma P_2$. ■

APPENDIX D

SET-VALUED MAPPINGS

Definition D.1 (Domain)**.** Given a set-valued mapping $F \colon \mathbb{R}^n \rightrightarrows \mathbb{R}^m$

$$\operatorname{dom} F = \{x \in \mathbb{R}^n \colon F(x) \neq \emptyset\}$$

Definition D.2 (Local boundedness)**.** A set-valued mapping $F \colon \mathbb{R}^n \rightrightarrows \mathbb{R}^m$ is locally bounded at x if there exists a neighborhood U_x of x such that $F(U_x)$ is bounded. F is locally bounded if it is locally bounded at each x. Given a set $S \subset \mathbb{R}^n$, F is locally bounded relatively to S if the set-valued mapping $\widehat{F} \colon \mathbb{R}^n \rightrightarrows \mathbb{R}^m$

$$\widehat{F}(x) \colon x \mapsto \begin{cases} F(x) & x \in S \\ \emptyset & x \notin S \end{cases}$$

is locally bounded at each $x \in S$.

Definition D.3 (Outer semicontinuity)**.** A set-valued mapping $F \colon \mathbb{R}^n \rightrightarrows \mathbb{R}^m$ is outer semicontinuous at $x \in \mathbb{R}^n$ if for every sequence of points x_k convergent to x and any convergent sequence of points $y_k \in F(x_k)$, one has $\lim y_k = y \in F(x)$. The mapping F is outer semicontinuous if it is outer semicontinuous at each $x \in \mathbb{R}^n$. Given a set $S \subset \mathbb{R}^n$, F is outer semicontinuous relatively to S if the set-valued mapping $\widehat{M} \colon \mathbb{R}^n \rightrightarrows \mathbb{R}^m$

$$\widehat{F}(x) \colon x \mapsto \begin{cases} F(x) & x \in S \\ \emptyset & x \notin S \end{cases}$$

outer semicontinuous at each $x \in S$.

This page is intentionally left blank.

GENERAL CONCLUSION AND RECOMMENDATIONS FOR FUTURE RESEARCH

In this dissertation, two specific problems arising in modern control systems were addressed. On the one hand, stability analysis and stabilization for quantized LTI continuous-time control systems. On the other hand, state estimation and observer-based control in the presence of both sporadic sensing and actuation for the case of LTI continuous-time systems. Although the two considered problems are tackled separately, the applicability of the results issued from our research situates in the context of control systems with limited information. Such a class of systems encompasses control systems built in the presence of communication constraints and/or in the presence of limited sensing and actuation capabilities.

The methodology offered within the first part of this thesis leads to constructive computer-aided tools for the analysis and the design of stabilizing controllers in the presence of actuator and sensor quantization. Both static state feedback controllers and dynamic output feedback controllers were considered, providing tools having a wide range of applications in real-world settings. Basically, given a LTI continuous-time plant subject to (uniform) quantization, either in the actuation channel or in the sensor channel or in both, the methodology we provided allows to design a LTI continuous-time controller ensuring uniform global asymptotic stability of a compact set containing the origin, while enabling the shrinkage of such a set via convex optimization.

The methodology offered within the second part of this thesis leads to constructive computer-aided tools for the design of asymptotic observers that exponentially reconstruct the state of a given LTI continuous-time plant, whenever the output is measured sporadically and only a lower and an upper bound on the sampling interval is known. Moreover, such observers can be used to stabilize a given LTI continuous-time plant in the presence of both sporadic sensing and actuation. Concerning, it was shown that the design of the resulting output feedback controller can be performed in two stages thanks to a separation principle.

Perspectives and Future Directions

As pointed throughout the conclusive chapters of this thesis, the work presented lets several questions open. In particular, within the scope of the first part, the extension of the offered methodology to other class of quantizers such as saturating quantizers is undoubtedly interesting and currently under investigation. Another interesting aspect, that is part of our current research, concerns the chattering suppression achieved by mean of hysteretic quantizers (see page 129 for further details). A likewise interesting aspect pertains to the development of alternative algorithms to handle the bilinear terms affecting the derived conditions. A worth improvement along that direction could be the derivation of more advanced strategies to improve the search of the optima, like in [96].

As far as concerns the second part, the main aspects to investigate pertain to the extension of the illustrated methodology to more general plant dynamics and to multi-ouput plants with asynchronous sampled channels; see page 219 for further details. In this setting, considering MIMO plants, another aspect to address is the construction of an observer-based controller to account both sporadic sensing and actuation, in the presence of asynchronous channels both in the input and in the output. This extension is worthwhile since it would help the design of output feedback controllers for networked control systems ([62]) via the use of a separation principle.

CV of the Author

Francesco Ferrante

Contact Information

3, Rue Bobillot
Toulouse, France 31200
http://homepages.laas.fr/fferante/

+33 (0)787761510
francesco.ferrante.2011@ieee.org

Languages

- Italian (native)
- English
- French

Research Interests

Quantization in control systems, nonlinear control, hybrid systems, observer design, limited information control systems, linear matrix inequalities in control.

Education

Institut Supérieur de l'Aéronautique et de l'Espace, Toulouse, France

Ph.D. Control engineering, *Expected:* Fall 2015

- Thesis Topic: *Control and estimation of dynamical systems in the presence of limited information*

Università degli Studi di Roma Tor Vergata, Rome, Italy

M.S., Control Engineering, Oct 2012

- Thesis: *On Quantization in Control Systems: Stabilization of Saturated Systems Subject to Quantization.*
- ***Summa Cum Laude***
- Advisors: Sergio Galeani, Sophie Tarbouriech and Frédéric Gouaisbaut

Università degli Studi di Roma Sapienza, Rome, Italy

B.S. Control Engineering, January 2010

- Thesis: *Development and experimentation of reactive motion strategies for the mobile robot Khepera III*
- Final score: ***110/110***
- Advisor: Giuseppe Oriolo

Research Experience

Research Assistant — October 2012 to present
Methods and algorithm in control,
Laboratoire d'analyse et d'architecture des systèmes,
Toulouse, France.
Supervisors: Sophie Tarbouriech, Frédéric Gouaisbaut

Visiting Scholar — July 2014 to August 2014
Computer Engineering Department,
University of California,
Santa Cruz, CA 95064, USA.
Supervisor: Ricardo G. Sanfelice

Intern — March 2012 to September 2012
Laboratoire d'analyse et d'architecture des systèmes,
Toulouse, France.
Supervisors: Sophie Tarbouriech, Frédéric Gouaisbaut

Research assistant (Bachelor level) — August to December 2009
Robotics Laboratory of the Dipartimento di Ingegneria Informatica, Automatica e Gestionale,
Sapienza University of Rome, Italy.
Supervisor: Giuseppe Oriolo

Journal Publications

1. F. Ferrante, F. Gouaisbaut, S. Tarbouriech, "Stabilization of Continuous-Time Linear Systems subject to Input Quantization", Automatica, 58(8):167-172, 2015.

Publications in refereed conference proceedings

1. F. Ferrante, F. Gouaisbaut, S. Tarbouriech, "Observer-based control for linear systems with quantized output", 13th European Control Conference, Strasbourg, France.
2. F. Ferrante, F. Gouaisbaut, R. G. Sanfelice, S. Tarbouriech, "An Observer with Measurement-triggered Jumps for Linear Systems with Known Input", 19th World Congress of the International Federation of Automatic Control, Cape Town, South Africa.
3. F. Ferrante, F. Gouaisbaut, R. G. Sanfelice, S. Tarbouriech, "Observer-based Control Design for Linear Systems in the Presence of Limited Measurement Streams and Intermittent Input Access", American Control Conference 2015, Chicago.
4. F. Ferrante, F. Gouaisbaut, S. Tarbouriech, "Dynamic Output-feedback Controller Design for Continuous-time Linear Systems with Actuator and Sensor Quantization", To appear in the proceedings of the 14th European Control Conference, Linz, Austria.
5. F. Ferrante, F. Gouaisbaut, R. G. Sanfelice, S. Tarbouriech, "A Hybrid Observer with a Continuous Intersample Injection in the Presence of Sporadic Measurements", To appear in the proceedings of the 54th IEEE Conference on Decision and Control, Osaka, Japan.

Book Chapters

1. F. Ferrante, F. Gouaisbaut, S. Tarbouriech, "Stabilization by quantized delayed state feedback", To appear in Delays in Networked Control Systems, in Advances in Delays and Dynamics series, Springer.

Unpublished (submitted and under review)

1. F. Ferrante, F. Gouaisbaut, R. G. Sanfelice, S. Tarbouriech, "State Estimation of Linear Systems in the Presence of Sporadic Measurements", Submitted to Automatica.

Papers currently in preparation

1. F. Ferrante, F. Gouaisbaut, S. Tarbouriech, "Stabilization of Nonlinear Systems subject to Input Hysteretic-quantization"
2. F. Ferrante, F. Gouaisbaut, S. Tarbouriech, "Stability and Stabilization of Linear Systems in the Presence of Sensor Quantization: An LMI-based approach"

Talks

- "Observer-based Control Design for Linear Systems in the Presence of Limited Measurement Streams and Intermittent Input Access". American Control Conference 2015 (ACC 2015), Chicago, Illinois (US), July 1-3, 2015.
- "An Observer with Measurement-Triggered Jumps for Linear Systems with Known Input". 19th World Congress of the International Federation of Automatic Control (IFAC 2014), Cape Town, South Africa, August 24-29, 2014.
- "Practical stabilization of linear delayed input quantized systems". 2nd Workshop on Delay Systems (Delsys 2013), Toulouse, France, November 20-22, 2013.

Academic community involvement

Technical Reviewer, Institute of Electrical and Electronic Engineers Inc.,
- IEEE Transaction on Automatic Control
- IEEE Conference on Decision and Control

Technical Reviewer,
- European Control Conference
- Journal Européen des Systèmes Automatisés

Teaching Experience

Teaching Assistant: Directly selected and enrolled by
Université Paul Sabatier, Toulouse, France — Fall 2012-present

Mentions include: Prepare practicals sessions, grade students, prepare integrative lecture notes.

- 2nd year undergraduate course on "Automatic control" (Practical sessions) Winter 2012, 2013, 2014.

- 3rd year undergraduate course on "Automatic control" (Practical sessions) Winter, Spring 2012, 2013, 2014.

- Graduate course on "Performance and robustness of feedback control systems" (Practical and tutorial sessions) — Fall 2013, 2014.

- Graduate course on "State space methods for linear control" (Practical sessions) Winter 2013

- Graduate course on "Nonlinear systems" (Co-lecturer) — Winter 2014

Professional memberships

- Institute of Electrical and Electronics Engineers (IEEE), Student Member
- IEEE Control Systems Society, Student Member
- IEEE Power Electronics Society, Student Member

Awards

- French Government, International exchange program fellowship — 2014
- French Government, Doctoral fellowship — 2012-2015
- European Commission, Student Placement fellowship — 2012
- University of Tor Vergata, Academic Excellence studentship — 2011-2012

References

Ricardo G. Sanfelice, PhD
E-mail: ricardo@ucsc.edu

Department of Computer Engineering,
University of California,
1156 High Street Santa Cruz, CA 95064, USA

Sophie Tarbouriech, PhD (PhD Supervisor)
E-mail: tarbour@laas.fr

Laboratoire d'analyse et d'architecture des systèmes
7, avenue du Colonel Roche,
31031 Toulouse, France

Frédéric Gouaisbaut, PhD (PhD Supervisor)

E-mail: fgouaisb@laas.fr

Université Paul Sabatier and
Laboratoire d'analyse et d'architecture des systèmes
7, avenue du Colonel Roche,
31031 Toulouse, France

Luca Zaccarian, PhD

E-mail: zaccarian@laas.fr

Università di Trento and
Laboratoire d'analyse et d'architecture des systèmes
7, avenue du Colonel Roche,
31031 Toulouse, France

BIBLIOGRAPHY

[1] M. Abbaszadeh and H. J. Marquez. “Robust $\mathcal{H}_\infty$ observer design for sampled-data Lipschitz nonlinear systems with exact and Euler approximate models”. *Automatica* 44.3 (2008), pp. 799–806.

[2] F. Amato, C. Cosentino, and A. Merola. “Stabilization of bilinear systems via linear state feedback control”. *Proceedings of the 15th Mediterranean Conference on Control & Automation*. 2007, pp. 1–5.

[3] V. Andrieu, M. Nadri, U. Serres, and J. Vivalda. “Continuous Discrete Observer with Updated Sampling Period”. *NOLCOS*. Vol. 9. 2013, pp. 439–444.

[4] M. Arcak and D. Nešić. “A framework for nonlinear sampled-data observer design via approximate discrete-time models and emulation”. *Automatica* 40.11 (2004), pp. 1931–1938.

[5] D. Arzelier and D. Peaucelle. “An iterative method for mixed $\mathcal{H}_2/\mathcal{H}_\infty$ synthesis via static output-feedback”. *Proceedings of the 41st IEEE Conference on Decision and Control*. Vol. 3. 2002, pp. 3464–3469.

[6] K. J. Aström and B. Wittenmark. *Computer-controlled systems: Theory and Design*. Courier Corporation, 2011.

[7] J.-P. Aubin and A. Cellina. *Differential inclusions*. Springer, 1984.

[8] A. Bacciotti. “Some remarks on generalized solutions of discontinuous differential equations”. *International Journal of Pure and Applied Mathematics* 10 (2004), pp. 257–266.

[9] A. Bacciotti. “Stabilization by means of state space depending switching rules”. *Systems & Control Letters* 53.3 (2004), pp. 195–201.

[10] A. Bacciotti and L. Rosier. “Liapunov and Lagrange stability: inverse theorems for discontinuous systems”. *Mathematics of Control, Signals and Systems* 11.2 (1998), pp. 101–128.

[11] I. Benedetti and E. A. Panasenko. “Positive invariance and differential inclusions with periodic right-hand side”. *Nonlinear Dynamics and System Theory* 7.4 (2007), pp. 339–349.

[12] G. M. Bergman. *An invitation to general algebra and universal constructions.* H. Helson, 1998.

[13] D. P. Bertsekas. *Convex optimization theory.* Athena Scientific Belmont, MA, 2009.

[14] S. Boyd and L. Vandenberghe. *Convex optimization.* Cambridge university press, 2004.

[15] S. Boyd, L. E. Ghaoui, E. Feron, and V. Balakrishnan. *Linear Matrix Inequalities in System and Control Theory.* Society for Industrial and Applied Mathematics, June 1997.

[16] A. Bressan. “Singularities of stabilizing feedbacks”. *Rend. Sem. Mat. Univ. Pol. Torino* 56.4 (1998), pp. 87–104.

[17] R. W. Brockett and D. Liberzon. “Quantized feedback stabilization of linear systems”. *IEEE Transactions on Automatic Control* 45.7 (2000), pp. 1279–1289.

[18] C. Cai and A. R. Teel. “Characterizations of input-to-state stability for hybrid systems”. *Systems & Control Letters* 58.1 (2009), pp. 47–53.

[19] D. Carnevale and A. R. Teel. “A Lyapunov proof of an improved maximum allowable transfer interval for networked control systems”. *IEEE Transactions on Automatic Control* 52.5 (2007), p. 892.

[20] E. B. Castelan, S. Tarbouriech, J. M. Gomes da Silva Jr., and I. Queinnec. “$\mathcal{L}_2$-stabilization of continuous-time linear systems with saturating actuators”. *Int. J. of Robust and Nonlinear Control* 16 (2006), pp. 935–944.

[21] F. Ceragioli and C. De Persis. “Discontinuous stabilization of nonlinear systems: Quantized and switching controls”. *Systems & control letters* 56.7 (2007), pp. 461–473.

[22] F. Ceragioli, C. De Persis, and P. Frasca. “Discontinuities and hysteresis in quantized average consensus”. *Automatica* 47 (2011), pp. 1919–1928.

[23] F. M. Ceragioli. “Discontinuous ordinary differential equations and stabilization”. Available online at http://calvino.polito.it/ ceragioli/tesidef.pdf. PhD thesis. Universtità degli studi di Firenze, Dipartimento di Matematica “Ulisse Dini”, 2000.

[24] G. Chesi. *Domain of attraction: analysis and control via SOS programming.* Springer Science & Business Media, 2011.

[25] M. Chilali and P. Gahinet. “$\mathcal{H}_\infty$ design with pole placement constraints: an LMI approach”. *IEEE Transactions on Automatic Control* 41.3 (1996), pp. 358–367.

[26] F. H. Clarke. “Discontinuous feedback and nonlinear systems”. *8th IFAC Symposium on Nonlinear Control Systems.* 2010, pp. 1–29.

[27] F. H. Clarke. *Nonsmooth analysis and control theory*. Vol. 178. Springer Science & Business Media, 1998.

[28] F. H. Clarke, Y. S. Ledyaev, and R. J. Stern. "Asymptotic stability and smooth Lyapunov functions". *Journal of differential Equations* 149.1 (1998), pp. 69–114.

[29] M. B. G. Cloosterman, L. Hetel, N. van de Wouw, W. P. M. H. Heemels, J. Daafouz, and H. Nijmeijer. "Controller synthesis for networked control systems". *Automatica* 46.10 (2010), pp. 1584–1594.

[30] J. B. Conway. *Functions of One Complex Variable II*. Springer Science & Business Media, 1995.

[31] J. Cortès. "Discontinuous dynamical systems – a tutorial on solutions, nonsmooth analysis, and stability". *IEEE Control Systems Magazine* 28.3 (2008), pp. 36–73.

[32] D. F. Coutinho, Minyue Fu, and C.E. de Souza. "Input and Output Quantized Feedback Linear Systems". *IEEE Transactions on Automatic Control* 55.3 (Mar. 2010), pp. 761 –766.

[33] J. Daafouz and J. Bernussou. "Parameter dependent Lyapunov functions for discrete time systems with time varying parametric uncertainties". *Systems & control letters* 43.5 (2001), pp. 355–359.

[34] D. F. Delchamps. "Extracting state information from a quantized output record". *Systems & Control Letters* 13.5 (1989), pp. 365–372.

[35] D. F. Delchamps. "Stabilizing a linear system with quantized state feedback". *IEEE Transactions on Automatic Control* 35.8 (Aug. 1990), pp. 916 –924.

[36] N. Elia and S. K. Mitter. "Stabilization of linear systems with limited information". *IEEE Transactions on Automatic Control* 46.9 (2001), pp. 1384–1400.

[37] F. Ferrante, F. Gouaisbaut, and S. Tarbouriech. "Dynamic Output-feedback Controller Design for Continuous-time Linear Systems with Actuator and Sensor Quantization". *To appear in the proceedings of the 14th European Control Conference*. 2015.

[38] F. Ferrante, F. Gouaisbaut, and S. Tarbouriech. "Observer-based control for linear systems with quantized output". *Proceedings of the 13th European Control Conference*. Updated version available online at https://hal.archives-ouvertes.fr/hal-00911570v2. 2014.

[39] F. Ferrante, F. Gouaisbaut, and S. Tarbouriech. "Stabilization by quantized delayed state feedback". *Delays in Networked Control Systems*. Ed. by A. Seuret, L. Hetel, J. Daafouz, and K. H. Johansson. Springer.

[40] F. Ferrante, F. Gouaisbaut, and S. Tarbouriech. "Stabilization of Continuous-Time Linear Systems subject to Input Quantization". *Automatica* 58 (2015), pp. 167–172.

[41] F. Ferrante, F. Gouaisbaut, R. G. Sanfelice, and S. Tarbouriech. "A Hybrdid Observer with a Continuous Intersample Injection in the Presence of Sporadic Measurements". *To appear in the proceedings of the 54th IEEE Conference on Decision and Control*. 2015.

[42] F. Ferrante, F. Gouaisbaut, R. G. Sanfelice, and S. Tarbouriech. "An Observer with Measurement-triggered Jumps for Linear Systems with Known Input". *Proceedings of the 19th World Congress of the International Federation of Automatic Control.* Vol. 19. 2014, pp. 140–145.

[43] F. Ferrante, F. Gouaisbaut, R. G. Sanfelice, and S. Tarbouriech. "Observer-based Control Design for Linear Systems in the Presence of Limited Measurement Streams and Intermittent Input Access". *Proceedings of the American Control Conference (ACC), 2015*, pp. 4689–4694.

[44] F. Ferrante, F. Gouaisbaut, R. G. Sanfelice, and S. Tarbouriech. "State Estimation of Linear Systems in the Presence of Sporadic Measurements". *Submitted at Automatica (currently under review)* (2015).

[45] F. Fichera. "Lyapunov techniques for a class of hybrid systems and reset controller syntheses for continuous-time plants". PhD thesis. ISAE-TOULOUSE, 2013.

[46] A. F. Filippov. *Differential equations with discontinuous right-hand side.* Dordrecht, The Netherlands: Kluwer, 1988.

[47] F. Forni, S. Galeani, L. Zaccarian, and D. Nešić. "Event-triggered transmission for linear control over communication channels". *automatica* 50.2 (2014), pp. 490–498.

[48] B. A. Francis and W. M. Wonham. "The internal model principle of control theory". *Automatica* 12.5 (1976), pp. 457–465.

[49] E. Fridman and M. Dambrine. "Control under quantization, saturation and delay: An LMI approach". *Automatica* 45 (2009), pp. 2258–2264.

[50] E. Fridman, U. Shaked, and V. Suplin. "Input/output delay approach to robust sampled-data $\mathcal{H}_\infty$ control". *Systems & Control Letters* 54.3 (2005), pp. 271–282.

[51] M. Fu and L. Xie. "The sector bound approach to quantized feedback control". *IEEE Transactions on Automatic Control* 50.11 (2005), pp. 1698–1711.

[52] P. Gahinet and P. Apkarian. "A linear matrix inequality approach to $\mathcal{H}_\infty$ control". *International Journal of Robust and Nonlinear Control* 4.4 (1994), pp. 421–448.

[53] G. Garcia, S. Tarbouriech, J.M. Gomes da Silva Jr., and D. Eckhard. "Finite $\mathcal{L}_2$ gain and internal stabilisation of linear systems subject to actuator and sensor saturations". *IET Control Theory Applications* 3.7 (July 2009), pp. 799–812.

[54] J. C. Geromel and M. C. de Oliveira. "H_2 and H_∞ robust filtering for convex bounded uncertain systems". *IEEE Transactions on Automatic Control* 46.1 (2001), pp. 100–107.

[55] R. Goebel, R. G. Sanfelice, and A. R. Teel. "Hybrid dynamical systems". *Control Systems Magazine* 29.2 (2009), pp. 28–93.

[56] R. Goebel, R. G. Sanfelice, and A. R. Teel. *Hybrid Dynamical Systems: Modeling, Stability, and Robustness.* Princeton University Press, 2012.

[57] R. Goebel and A. R. Teel. "Preasymptotic stability and homogeneous approximations of hybrid dynamical systems". *SIAM review* 52.1 (2010), pp. 87–109.

[58] M. Green and D. JN. Limebeer. *Linear robust control*. Courier Dover Publications, 2012.

[59] O. Hájek. "Discontinuous differential equations, I". *Journal of Differential Equations* 32.2 (1979), pp. 149–170.

[60] W. P. M. H. Heemels, N. van de Wouw, R. H. Gielen, M. C. F. Donkers, L. Hetel, S. Olaru, M. Lazar, J. Daafouz, and S. Niculescu. "Comparison of overapproximation methods for stability analysis of networked control systems". *Proceedings of the 13th ACM International Conference on Hybrid Systems: Computation and Control*. ACM. 2010, pp. 181–190.

[61] D. Henrion and A. Garulli. *Positive polynomials in control*. Vol. 312. Springer Science & Business Media, 2005.

[62] J. P. Hespanha, P. Naghshtabrizi, and Y. Xu. "A survey of recent results in networked control systems". *Proceeding of the IEEE* 95.1 (2007), p. 138.

[63] L. Hetel, J. Daafouz, and C. Iung. "Stabilization of Arbitrary Switched Linear Systems With Unknown Time-Varying Delays". *IEEE Transactions on Automatic Control,* 51.10 (Oct. 2006), pp. 1668–1674.

[64] L. Hetel, J. Daafouz, S. Tarbouriech, and C. Prieur. "Stabilization of linear impulsive systems through a nearly-periodic reset". *Nonlinear Analysis: Hybrid Systems* 7.1 (2013), pp. 4–15.

[65] D. Hristu-Varsakelis and W. S. Levine. *Handbook of networked and embedded control systems*. Springer, 2005.

[66] T. Hu and Z. Lin. *Control systems with actuator saturation: Analysis and design*. Boston: Birkhauser, 2001.

[67] Quanser Consulting Inc. *Rotary Inverted Pendulum ROTPEN*. 1998.

[68] A. Isidori and C. I. Byrnes. "Output regulation of nonlinear systems". *IEEE Transactions on Automatic Control* 35.2 (1990), pp. 131–140.

[69] T. Kailath. *Linear systems*. Vol. 1. Prentice-Hall Englewood Cliffs, NJ, 1980.

[70] R. E. Kalman. "Nonlinear aspects of sampled-data control systems." *Proceedings of the symposium on nonlinear circuit theory Brooklyn, NY*. 1956.

[71] P. Kapasouris, M. Athans, and G. Stein. "Design of feedback control systems for stable plants with saturating actuators". *Proceedings of the 27th IEEE Conference on Decision and Control*. 1988, pp. 469–479.

[72] I. Karafyllis and C. Kravaris. "From Continuous-Time Design to Sampled-Data Design of Nonlinear Observers". *arXiv:0801.4824* (2008).

[73] I. Karafyllis and C. Kravaris. "From continuous-time design to sampled-data design of observers". *IEEE Transactions on Automatic Control,* 54.9 (2009), pp. 2169–2174.

[74] N. Kazantzis and C. Kravaris. "Discrete-time nonlinear observer design using functional equations". *Systems & Control Letters* 42.2 (2001), pp. 81–94.

[75] C. M. Kellett and A. R. Teel. "Smooth Lyapunov functions and robustness of stability for difference inclusions". *Systems & Control Letters* 52.5 (2004), pp. 395–405.

[76] H. K. Khalil. *Nonlinear Systems*. 3rd. Prentice-Hall, 2002.

[77] N. Krasovskii. *Stability of motion: applications of Lyapunov's second method to differential systems and equations with delay*. Stanford university press, 1963.

[78] M. Lazar, W.P. M. H. Heemels, and A. R. Teel. "Lyapunov functions, stability and input-to-state stability subtleties for discrete-time discontinuous systems". *IEEE Transactions on Automatic Control* 54.10 (2009), pp. 2421–2425.

[79] H. Lens and J. Adamy. "Observer Based Controller Design for Linear Systems with Input Constraints". *IFAC World Congress*. Vol. 17. 1. 2008, pp. 9916–9921.

[80] J. Leyva-Ramos. "A new look at partial fraction expansion of transfer function matrices from a computational viewpoint". *Computers & Mathematics with Applications* 26.3 (1993), pp. 27–35.

[81] Y. Li and R. G. Sanfelice. "A Robust Finite-time Convergent Hybrid Observer for Linear Systems". *Proceedings of the 52th IEEE Conference on Decision and Control*. 2013.

[82] D. Liberzon. "Hybrid feedback stabilization of systems with quantized signals". *Automatica* 39 (2003), pp. 1543–1554.

[83] D. Liberzon. "Observer-based quantized output feedback control of nonlinear systems". *Mediterranean Conference on Control Automation, 2007. MED '07.* (June 2007), pp. 1–5.

[84] D. Liberzon. *Switching in Systems and Control*. Boston, USA: Birkhauser, 2003.

[85] H. Liu, M. Cao, and C. De Persis. "Quantization effects on synchronized motion of teams of mobile agents with second-order dynamics". *Systems & Control Letters* 61.12 (2012), pp. 1157–1167.

[86] J. Liu and A. R. Teel. "Hybrid Systems with Memory: Modelling and Stability Analysis via Generalized Solutions". *Proceedings of the 19th World Congress of the International Federation of Automatic Control*. Vol. 19. 1. 2014, pp. 6019–6024.

[87] J. Lofberg. "YALMIP : a toolbox for modeling and optimization in MATLAB". *IEEE International Symposium on Computer Aided Control Systems Design, 2004*. Sept. 2004, pp. 284–289.

[88] D. G. Luenberger. "An introduction to observers". *IEEE Transactions on Automatic Control* 16.6 (1971).

[89] R. K. Miller, M. S. Mousa, and A. N. Michel. "Quantization and overflow effects in digital implementations of linear dynamic controllers". *IEEE Transactions on Automatic Control,* 33.7 (July 1988), pp. 698–704.

[90] M. Moarref and L. Rodrigues. “Observer design for linear multi-rate sampled-data systems”. *American Control Conference 2014*. June 2014, pp. 5319–5324.

[91] M. Moarref and L. Rodrigues. “Stability and stabilization of linear sampled-data systems with multi-rate samplers and time driven zero order holds”. *Automatica* 50.10 (2014), pp. 2685–2691.

[92] P. E. Moraal and J. W. Grizzle. “Observer design for nonlinear systems with discrete-time measurements”. *IEEE Transactions on Automatic Control* 40.3 (1995), pp. 395–404.

[93] R. M. Murray, K. Aström, S. Boyd, R. Brockett, and G. Stein. “Control in an information rich world”. *Report of the Panel on Future Directions in Control Theory* (2002).

[94] M. Nadri and H. Hammouri. “Design of a continuous-discrete observer for state affine systems”. *Applied Mathematics Letters* 16.6 (2003), pp. 967–974.

[95] D. Nešić, A. R. Teel, G. Valmorbida, and L. Zaccarian. “Finite-gain $\mathcal{L}_p$ stability for hybrid dynamical systems”. *Automatica* 49.8 (2013), pp. 2384–2396.

[96] E. Ostertag. “An Improved Path-Following Method for Mixed Controller Design”. *IEEE Transactions on Automatic Control* 53.8 (2008), pp. 1967–1971.

[97] B. Paden and S. Sastry. “A calculus for computing Filippov’s differential inclusion with application to the variable structure control of robot manipulators”. *IEEE Transactions on Circuits and Systems* 34.1 (1987), pp. 73–82.

[98] G. Peano. “Démonstration de l’intégrabilité des équations différentielles ordinaires”. *Mathematische Annalen* 37.2 (1890), pp. 182–228.

[99] G. Pipeleers, B. Demeulenaere, J. Swevers, and L. Vandenberghe. “Extended LMI characterizations for stability and performance of linear systems”. *Systems & Control Letters* 58.7 (2009), pp. 510–518.

[100] R. Postoyan and D. Nešić. “A framework for the observer design for networked control systems”. *IEEE Transactions on Automatic Control* 57.5 (2012), pp. 1309–1314.

[101] R. Postoyan, P. Tabuada, D. Nešić, and A. Anta. “Event-triggered and self-triggered stabilization of distributed networked control systems”. *Proceedings of the 50th IEEE Conference on Decision and Control and European Control Conference 2011 (CDC-ECC)*. 2011, pp. 2565–2570.

[102] A. S. Poznyak. *Advanced Mathematical Tools for Automatic Control Engineers, 1*. Elsevier, 2008.

[103] T. Raff and F. Allgöwer. “Observers with impulsive dynamical behavior for linear and nonlinear continuous-time systems”. *Proceedings of the 46th IEEE Conference on Decision and Control*. 2007, pp. 4287–4292.

[104] T. Raff, M. Kogel, and F. Allgöwer. “Observer with sample-and-hold updating for Lipschitz nonlinear systems with nonuniformly sampled measurements”. *Proceedings of the American Control Conference, 2008*. IEEE. 2008, pp. 5254–5257.

[105] M. Robles and R. G. Sanfelice. "Hybrid controllers for tracking of impulsive reference state trajectories: a hybrid exosystem approach". *Proceedings of the 14th international conference on Hybrid systems: Computation and Control*. ACM. 2011.

[106] R. T. Rockafellar and R. J-B. Wets. *Variational analysis*. Vol. 317. Springer Science & Business Media, 2009.

[107] W. Rudin. *Principles of mathematical analysis*. Vol. 3. McGraw-Hill New York, 1964.

[108] R. G. Sanfelice, D. Copp, and P. Nanez. "A toolbox for simulation of hybrid systems in Matlab/Simulink: Hybrid Equations (HyEQ) Toolbox". *Proceedings of the 16th international conference on Hybrid Systems: Computation and Control*. ACM. 2013, pp. 101–106.

[109] R. G. Sanfelice and L. Praly. "Convergence of nonlinear observers on $\mathbb{R}^n$ with a Riemannian metric (Part I)". *IEEE Transactions on Automatic Control* 57.7 (2012), p. 1709.

[110] C. W. Scherer. "The state-feedback $\mathcal{H}_\infty$-problem at optimality". *Automatica* 30.2 (1994), pp. 293–305.

[111] R. Sentis. "Equations différentielles à second membre mesurable". *Boll. Un. Mat. Ital* 15 (1978), pp. 724–742.

[112] B. Sinopoli, L. Schenato, M. Franceschetti, K. Poolla, M. I. Jordan, and S.S. Sastry. "Kalman filtering with intermittent observations". *IEEE Transactions on Automatic Control,* 49.9 (2004), pp. 1453–1464.

[113] B. Sinopoli, L. Schenato, M. Franceschetti, K. Poolla, and S. S. Sastry. "Optimal control with unreliable communication: the TCP case". *Proceedings of the American Control Conference*. Vol. 5. June 2005, pp. 3354–3359.

[114] E. D. Sontag. "Input to state stability: basic concepts and results". *Nonlinear and Optimal Control Theory*. Ed. by P. Nistri and G. Stefani. Berlin, Germany: Springer-Verlag, 2007, pp. 163–220.

[115] E. D. Sontag. "Smooth stabilization implies coprime factorization". *IEEE Transactions on Automatic Control* 34.4 (1989), pp. 435–443.

[116] J. Sur and B. Paden. "Observers for linear systems with quantized outputs". *Proceedings of the American Control Conference*. Vol. 5. 1997, pp. 3012 –3016.

[117] S. Tarbouriech and F. Gouaisbaut. "Control Design for Quantized Linear Systems With Saturations". *IEEE Transactions on Automatic Control* 57.7 (July 2012), pp. 1883 –1889.

[118] S. Tarbouriech, I. Queinnec, and C. Prieur. "Stability analysis and stabilization of systems with input backlash". *IEEE Transactions on Automatic Control* 59.2 (2014), pp. 488–494.

[119] S. Tarbouriech, I. Queinnec, and C. Prieur. "Stability analysis for systems with saturation and backlash in the loop". *Proceedings of the 52nd IEEE Conference on Decision and Control*. Dec. 2013, pp. 6652–6657.

[120] S. Tarbouriech, G. Garcia, J. M. Gomes da Silva Jr., and I. Queinnec. *Stability and stabilization of linear systems with saturating actuators.* Springer, 2011.

[121] S. Tatikonda and S. Mitter. “Control under communication constraints”. *IEEE Transactions on Automatic Control,* 49.7 (2004), pp. 1056–1068.

[122] A. R. Teel. “Observer-based hybrid feedback: a local separation principle”. *American Control Conference (ACC), 2010*, pp. 898–903.

[123] A. R. Teel, F. Forni, and L. Zaccarian. “Lyapunov-based sufficient conditions for exponential stability in hybrid systems”. *IEEE Transactions on Automatic Control* 58.6 (2013), pp. 1591–1596.

[124] A. R. Teel and L. Praly. “A smooth Lyapunov function from a class-$\mathcal{KL}$ estimate involving two positive semidefinite functions”. *ESAIM: Control, Optimisation and Calculus of Variations* 5 (2000), pp. 313–367.

[125] M. L. J. Van de Vel. *Theory of convex structures.* Vol. 50. Elsevier, 1993.

[126] J. G. VanAntwerp and R. D. Braatz. “A tutorial on linear and bilinear matrix inequalities”. *Journal of Process Control* 10.4 (2000), pp. 363–385.

[127] L. Vandenberghe and S. Boyd. “A polynomial-time algorithm for determining quadratic Lyapunov functions for nonlinear systems”. *European conference on circuit theory and design.* 1993, pp. 1065–1068.

[128] M. Vidyasagar. “On the stabilization of nonlinear systems using state detection”. *IEEE Transactions on Automatic Control* 25.3 (1980), pp. 504–509.

[129] G. C. Walsh, Hong Y., and L. G. Bushnell. “Stability analysis of networked control systems”. *IEEE Transactions on Control Systems Technology* 10.3 (May 2002), pp. 438–446.

This page is intentionally left blank.

LIST OF FIGURES

LIST OF TABLES

This page is intentionally left blank.

LIST OF SYMBOLS

$\mathbb{N}$	The set of strictly positive integer numbers
$\mathbb{N}_0$	The set of nonnegative integer numbers
$\mathbb{C}$	The set of complex numbers
$\mathbb{R}$	The set of real numbers
$\mathbb{R}_{\geq 0}$	The set of nonnegative real numbers
$\mathbb{R}_{>0}$	The set of strictly positive real numbers
$\mathbb{R}^{n\times m}$	The real $n \times m$ matrices space
$\mathbb{C}^{n\times m}$	The complex $n \times m$ matrices space
$\mathbb{R}^n$	The n-dimensional Euclidean space
$\mathbb{R}^n_+$	The subset of the n-dimensional Euclidean space composed by vectors with strictly positive components
$\lvert x\rvert$	The absolute value if $x \in \mathbb{R}$, the vector component-wise absolute value if $x \in \mathbb{R}^n$
$\lVert x\rVert$ ($\lVert A\rVert$)	The Euclidean norm of a vector x (The Euclidean induced matrix norm of a matrix A)
A^T	The transpose of a matrix A
$A^{-\mathsf{T}}$	Indicates $(A^{-1})^\mathsf{T}$ for any nonsingular matrix A
$A^{\frac{1}{2}}$	Indicates the square root of a symmetric positive definite matrix A, *i.e.*, $A^{\frac{1}{2}}A^{\frac{1}{2}} = A$
$A^{-\frac{1}{2}}$	Indicates the inverse square root of a symmetric positive definite matrix A, *i.e.*, $A^{-\frac{1}{2}}A^{-\frac{1}{2}} = A^{-1}$
(x, y)	Equivalent notation for the vector $[x^\mathsf{T}, y^\mathsf{T}]^\mathsf{T}$
$\langle x, y\rangle$	With $x, y \in \mathbb{R}^n$, stands for the standard Euclidean inner product
$\nabla f(x)$	With $f\colon \mathbb{R}^n \to \mathbb{R}$, stands for the gradient of f
$\mathcal{S}^n_+$	The set of $n \times n$ symmetric positive definite matrices
$\mathcal{D}^n_+$	The set of $n \times n$ diagonal positive definite matrices

$\mathbb{B}$	The closed unit ball, of appropriate dimension, in the Euclidean norm
$\mathrm{trace}(A)$	The trace of the matrix A
$\mathrm{He}(A)$	Stands for $A + A^{\mathsf{T}}$
$\mathrm{spec}(A)$	The spectrum of the matrix A
$\bullet$	Stands for symmetric blocks in block matrices
I_n	The $n \times n$ identity matrix
$\mathbf{0}$	the null scalar or the null matrix of appropriate dimension
$\mathbf{1}_n$	unitary vector of $\mathbb{R}^n$, *i.e.*, $\mathbf{1}_n = \underbrace{[1, 1, \ldots, 1]}_{n}$
$\mathcal{E}(A)$	With $A \in \mathcal{S}^n_+$, stands for the ellipsoidal set defined by $\{x \in \mathbb{R}^n \colon x'Ax \leq 1\}$
$x_{(i)}$	With $x \in \mathbb{R}^n$, stands for the i^{th} component of x
$\lambda_{\min}(A)(\lambda_{\max}(A))$	The smallest (largest) eigenvalue of a matrix $A \in \mathcal{S}^n_+$
$A > \mathbf{0}$	Means that A is positive definite
$A \geq \mathbf{0}$	Means that A is positive semi-definite
$A \geq B$	Means that the matrix $A - B \geq \mathbf{0}$
$A > B$	Means that the matrix $A - B > \mathbf{0}$
$x \prec y$	With $x, y \in \mathbb{R}^n$ means $x_{(i)} < y_{(i)}$ for each $i = 1, 2, \ldots, n$
$\mathrm{diag}(A_1, A_2, \ldots, A_s)$	Denotes the block-diagonal matrix whose diagonal elements are $A_1, A_2, \ldots, A_s$
$\dot{x}$	The derivative, with respect to time, of the state x of a continuous-time or hybrid dynamical system
$\mathrm{co}\, S$	The convex hull of the set S
$\overline{S}$	The closure of the set S
∂S	The boundary of the set S
$S_1 \times S_2$	Stands for the standard Cartesian product of the sets S_1, S_2.
$\bigtimes_{i=1}^{p} S_i$	Stands for the standard Cartesian product of the sets $S_1, S_2, \ldots, S_p$.
$\mathrm{Int} S$	The interior of the set S
$\overline{\mathrm{co}} S$	The closed convex hull of the set S
$\setminus$	Set difference
$\mathbf{card}(S)$	The cardinality of the set S
$\Re(x)$	The real part of $x \in \mathbb{C}$
$\Im(x)$	The imaginary part of $x \in \mathbb{C}$
$x^\star$	The complex-conjugate of $x \in \mathbb{C}$
$\Re(A)$	The real part of $A \in \mathbb{C}^m \times \mathbb{C}^n$
$\Im(A)$	The imaginary part of $A \in \mathbb{C}^m \times \mathbb{C}^n$
$A^\star$	The element-wise complex-conjugate of $A \in \mathbb{C}^m \times \mathbb{C}^n$

id	The identity function
$F\colon \mathbb{R}^n \rightrightarrows \mathbb{R}^m$	This notation indicates that F is a set-valued mapping with $F(x) \subset \mathbb{R}^m$ for each $x \in \mathbb{R}^n$